

FONDO EDITORIAL

CONGRESO INTERNACIONAL DE INGENIERÍA DE SISTEMAS



UNIVERSIDAD  
DE LIMA

# INNOVANDO LA EDUCACIÓN EN TECNOLOGÍA

ACTAS DEL II CONGRESO INTERNACIONAL  
DE INGENIERÍA DE SISTEMAS



*Congreso Internacional de Ingeniería de Sistemas*

---

Innovando la educación en tecnología  
Actas del II Congreso Internacional de Ingeniería de Sistemas



---

FONDO EDITORIAL

CONGRESO INTERNACIONAL DE INGENIERÍA DE SISTEMAS

---



UNIVERSIDAD  
DE LIMA

# INNOVANDO LA EDUCACIÓN EN TECNOLOGÍA

ACTAS DEL II CONGRESO INTERNACIONAL  
DE INGENIERÍA DE SISTEMAS

*Innovando la educación en tecnología. Actas del II Congreso Internacional de Ingeniería de Sistemas* / presentación, Nadia Rodríguez Rodríguez. Primera edición. Lima: Universidad de Lima, Fondo Editorial, 2020.

281 páginas: ilustraciones, gráficos, diagramas.

Texto en inglés y español.

Conferencias magistrales y ponencias del II Congreso Internacional de Ingeniería de Sistemas (CIIS), denominado "Innovando la educación en tecnología", el cual se realizó en la Universidad de Lima, los días 5 y 6 de septiembre de 2019.

Incluye referencias. Pósteres: páginas 251-281.

1. Ingeniería de *software* -- Congresos. 2. Innovaciones educativas -- Congresos.  
3. Tecnología educativa -- Congresos. I. Rodríguez-Rodríguez, Nadia-Katherine, presentación. II. Universidad de Lima. Fondo Editorial. III. Congreso Internacional de Ingeniería de Sistemas (2.º: 2019: Lima, Perú).

005.1

C

2

ISBN 978-9972-45-531-5

*Innovando la educación en tecnología. Actas del II Congreso Internacional de Ingeniería de Sistemas*  
Primera edición digital: julio, 2020

© De esta edición:  
Universidad de Lima  
Fondo Editorial  
Av. Javier Prado Este 4600,  
Urb. Fundo Monterrico Chico, Lima 33  
Apartado postal 852, Lima 100, Perú  
Teléfono: 437-6767, anexo 30131  
fondoeditorial@ulima.edu.pe  
www.ulima.edu.pe

Diseño, edición y carátula: Fondo Editorial de la Universidad de Lima

Prohibida la reproducción total o parcial de este libro, por cualquier medio,  
sin permiso expreso del Fondo Editorial.

ISBN 978-9972-45-531-5

## **PROGRAM COMMITTEE / COMITÉ ACADÉMICO**

- Juan Gutiérrez
- Manuela Linares
- Andrea Matuk
- Nadia Rodríguez

## **ORGANIZING COMMITTEE / COMITÉ ORGANIZADOR**

- Emma Barreno
- Lucinda Butrón
- Daniel Cárdenas
- Juan Gutiérrez
- Rosario Guzmán
- Andrea Matuk
- Rosa Millones
- Nadia Rodríguez
- Pablo Rojas
- Carlos Torres





# Índice

PRESENTACIÓN	11
CONFERENCIAS MAGISTRALES	15
Improvisation for Technically-Oriented People Russell McMahon	17
Gamificación en la educación superior Elvira Rincón-Flores	25
What can innovation in engineering education do for you as a student and what can you do as a student for innovation in engineering education? Carlos Alario-Hoyos	37
Building Blocks for Powerful Ideas: Designing a Programming Language to Teach the Beauty and Joy of Computing Jens Möning	45
PONENCIAS	47
Análisis de sentimientos de noticias escritas usando un modelo basado en la red neuronal <i>long short-term memory</i> para determinar si las noticias positivas mejoran el estado de ánimo de las personas Gustavo Adolfo Reyes-Paredes	49
Modelo de selección de plataforma educativa virtual con mapas cognitivos difusos (FCM) Manuela Linares-Barbero	63
Evaluación de un sistema de búsqueda de rutas de evacuación eficientes de un establecimiento usando el algoritmo D estrella Walter Steven Pariona-Sánchez	77

Detección de intrusiones basada en modelado de red resistente a evasión por técnicas de imitación Jorge Maestre-Vidal, Marco Antonio Sotelo-Monge	91
La estructura organizacional, el <i>agile mindset</i> y el gobierno de TI para la transformación digital de las empresas Marco A. Fernández del Pomar	107
Sistema de detección y clasificación de peces utilizando visión computacional Rodrigo Mejía, Gianfranco Rosales	127
Sistema para incremento de vocabulario para la comprensión lectora en primaria con ayuda de realidad aumentada Gianella Urday	143
Metrics to Understand Future Maintenance Effort Required of Complicated Source Code Michael Dorin, Sergio Montenegro	171
Red neuronal convolucional para la detección de aves exóticas en peligro de extinción Guillermo Eduardo Narváz	185
Aprender a programar sin el uso de la computadora Marjorie Tineo-Tineo	201
Propuesta de un patrón de arquitecturas de <i>software</i> para la interoperabilidad en dispositivos en la capa al borde de un ecosistema IoT Juan Moreno-Motta, Felipe Moreno-Vera, Frank Moreno-Vera	213
Generación de reglas de asociación para productos de <i>retail</i> utilizando el algoritmo FP-Growth paralelo Renato Pérez-Gómez	231
PÓSTERES	251

# Presentación

Este libro contiene un conjunto de conferencias magistrales y ponencias presentadas durante el II Congreso Internacional de Ingeniería de Sistemas (CIIS), denominado “Innovando la educación en tecnología”, el cual se realizó en la Universidad de Lima, los días 5 y 6 de septiembre del 2019. Este congreso fue dirigido principalmente a la comunidad académica nacional y extranjera y contó con la presencia de académicos internacionales reconocidos, docentes y estudiantes de posgrado y pregrado.

El CIIS 2019 tuvo como finalidad promover, estimular y difundir la investigación e innovación de la educación en carreras afines a la tecnología. La temática relacionada a la educación fue seleccionada, porque pensamos que educar es un reto y en tecnología es aún más. Asimismo, existen muchos desafíos por delante para lograr que el aprendizaje en las nuevas generaciones sea cada vez más atractivo y motivador.

Importantes especialistas internacionales pronunciaron conferencias magistrales y expusieron sus investigaciones. Russell McMahon, docente investigador en técnicas de improvisación para la enseñanza de tecnología, miembro de la School of Information Technology en la Universidad de Cincinnati (Estados Unidos), presenta un trabajo sobre técnicas teátricas de improvisación para mejorar la comunicación y las habilidades de formación, así como su uso en la enseñanza de conceptos técnicos como la ciberseguridad, diseño de base de datos, conceptos de programación y, lo más importante, cómo mejorar las habilidades de comunicación que permitan ayudar a los nuevos profesionales a afrontar los nuevos desafíos de su generación y de las próximas. El World Economic Forum ha publicado las habilidades necesarias para tener éxito en lo profesional para el 2020, la gran mayoría de ellas están relacionadas con las habilidades blandas, por lo que es sumamente interesante que McMahon haya descubierto en la técnica de “impro” una manera de hacer que los ingenieros nos comuniquemos efectivamente entre nosotros y con profesionales de otras especialidades.

La doctora Elvira G. Rincón Flores, docente en matemáticas y especialista del Tecnológico de Monterrey (México), presenta su investigación basada en la gamificación adoptada por los educadores de diversas partes del mundo con el propósito de enganchar y motivar a los estudiantes en su proceso de aprendizaje. La gamificación en la educación puede ser una herramienta valiosa que ayuda a la formación de los estudiantes trabajando en tres aspectos importantes como lo cognitivo, lo social y lo emotivo; mostrando resultados notablemente favorables. El investigador en neurociencia Francisco Mora asegura que el elemento esencial en el proceso de aprendizaje es la emoción, porque solo se puede aprender aquello que se ama, aquello que le dice algo nuevo a la persona que significa algo, que sale de lo monótono. Es justamente lo que Elvira argumenta, que los juegos provocan emociones y, en consecuencia, pueden ayudar a hacer el aprendizaje más efectivo.

El doctor Carlos Alario-Hoyos, profesor de la Universidad Carlos III de Madrid (España) y subdirector de la Cátedra Unesco de Educación Digital Escalable para Todos, nos muestra sus habilidades a través de la investigación y desarrollo en tecnología educativa, aprendizaje colaborativo y evaluación de experiencias de aprendizaje. Alario-Hoyos discute sobre el impulso de la innovación como factor importante para aplicar en la enseñanza de la ingeniería; asimismo el compromiso de los estudiantes con la educación. Enseña que es posible recopilar datos de *stakeholders* con mayor detalle; sin embargo, esta información recopilada debe transformarse en indicadores y visualizaciones de alto nivel que guíen los procesos de toma de decisiones por parte de los gestores académicos. Su investigación está orientada a evaluar cómo la innovación basada en datos que genera la educación de ingeniería permitirá mejoras continuas en el contenido educativo, repercutirá en las interacciones sociales de los estudiantes, las habilidades de aprendizaje mejoradas por los resultados analizados; así como un análisis de los posibles resultados.

Finalmente, Jens Mönig, investigador en SAP (Alemania), experto en entornos de programación interactiva y cocreador de la herramienta Snap!, nos muestra este lenguaje de programación bajo un entorno de programación gráfica nativo en la nube y dentro de una comunidad en línea. Snap! es utilizado en el plan de estudios Beauty and Joy of Computing en la Universidad de Berkeley (Estados Unidos) y empieza a propagarse en otras universidades. Asimismo, se imparte en colegios y escuelas secundarias para la enseñanza y la investigación y ha sido diseñado para su inclusión. Los niveles introductorios de la herramienta están desarrollados para estudiantes principiantes de todos los campos de estudio, no solo de carreras relacionadas con la computación, y sus funcionalidades multimedia invitan a ser utilizadas por pensadores creativos de todas las edades.

En este libro también se muestran los artículos seleccionados producto de nuestra convocatoria *call for papers*, la cual, como en nuestro primer CIIS, fue bastante exitosa. De igual manera, el libro muestra las sumillas de los pósteres académicos exhibidos en el CIIS. La convocatoria de *call for posters* fue la novedad en esta edición y tuvo por objetivo brindar a

autores y participantes la oportunidad de compartir sus propuestas de investigación finalizadas o en proceso, las experiencias y los desafíos experimentados en la integración de la tecnología en el proceso de enseñanza-aprendizaje de manera innovadora.

Cabe señalar que los artículos presentados, tanto por los investigadores nacionales como internacionales, fueron sometidos a un proceso de revisión por pares, el cual garantiza la calidad de los trabajos expuestos en el congreso. Los temas correspondieron a áreas de investigación definidas dentro del marco del CIIS: sistemas de información, ingeniería de *software*, tecnologías de la información y ciencias de la computación.

El congreso ha cumplido con ofrecer un espacio y la infraestructura necesaria para promover y difundir las investigaciones presentadas, logrando un ambiente participativo a través de escuchar diferentes puntos de investigación y permitiendo debates; lo cual ha promovido reflexionar sobre cómo la tecnología de la información puede aplicarse en diferentes campos generando grandes aportes. Asimismo, ha permitido la dinamización de las acciones en pro de la investigación desde los estudiantes de pregrado. Esperamos seguir generando valor agregado con todas las investigaciones mostradas en el congreso.

Agradecemos a las autoridades de la Universidad de Lima por su respaldo para la realización del CIIS 2019, así como al comité organizador, al personal administrativo y a los alumnos de la Carrera de Ingeniería de Sistemas, por brindarnos su valiosa colaboración, que hizo posible el desarrollo del evento, cumpliendo así con el objetivo de integración y fomento de la investigación.

*Nadia Rodríguez Rodríguez*  
Directora  
Carrera de Ingeniería de Sistemas  
Universidad de Lima



# CONFERENCIAS MAGISTRALES

---





# Improvisation for Technically-Oriented Peoples

Russell McMahon  
mcmahor@ucmail.uc.edu / University of Cincinnati, USA

Recepción: 17-7-2019 / Aceptación: 7-8-2019

**ABSTRACT.** Teaching “soft” skills to technical people is just as important as learning “hard” skills. Improvisation techniques can also be used in teaching technical concepts such as cybersecurity, agile development, database design, programming concepts, and most importantly how to better one’s communication skills. In an age where rapid changes have become the norm, improvisation techniques can be used to help navigate the new challenges of the next generation careers, global interaction, and technologies. These techniques can easily be incorporated in other methodologies such as creative problem-solving and design thinking. There are clearly defined and flexible rules for improvising, which make it easier for technical persons to learn and use in their daily life and career.

**KEYWORDS:** improvisation, innovation, creative problem-solving, technical education

## Improvisación para personas con orientación técnica

**RESUMEN.** Enseñar las habilidades “blandas” a las personas técnicas es tan importante como aprender las habilidades “duras”. Las técnicas de improvisación también se pueden usar en la enseñanza de conceptos técnicos como la ciberseguridad, el desarrollo ágil, el diseño de bases de datos, los conceptos de programación y, lo más importante, cómo mejorar las habilidades de comunicación. En una era en la que los cambios rápidos se han convertido en la norma, las técnicas de improvisación pueden usarse para ayudar a navegar los nuevos desafíos de las carreras de la próxima generación, la interacción global y las tecnologías. Estas técnicas pueden incorporarse fácilmente en otras metodologías, como la resolución creativa de problemas y el pensamiento de diseño. Existen reglas claramente definidas y flexibles para la improvisación que facilitan que las personas técnicas aprendan y usen en su vida diaria y carrera.

**PALABRAS CLAVE:** improvisación, innovación, resolución creativa de problemas, educación técnica

## 1. INTRODUCTION

The first known use of improvisation goes back to the Atellan farce (*fabula Atellana*) dating back as far as 391 BCE, but it seems to have faded by 55 BCE according to Cicero [12] and no written records appear to exist past the first century CE [7]. In the 1500s, the improvisation art form was revived through the *commedia dell'arte*. Although it was performed in an unscripted format, the characters, plot, and locale were pre-set. It lasted until about 1800 and died out, in part, due to the political nature of the shows themselves, which led to suppression and eventual outright banning.

Improvisation was resurrected as a theater art form in the 20th century through the independent work of Viola Spolin and Keith Johnston. Spolin defined what improvisation theater in the US would look like through her book “Improvisation for the Theater”, which was first published in 1963. Johnston wrote “Impro: Improvisation and the Theatre” and created Theatresports. It is through the work of these two that improvisation was codified and advanced in such a way that it could be applied in many STEM (Science, Technology, Engineering, Math)-related fields.

I started taking formal improvisation classes in January 2017 from a local improv company and have continued to do so even today. In September 2017, I began to give improv talks/workshops at mostly academic-related conferences. Over the next year, I spoke at a total of seven national/international conferences. In October 2017, I attended the Midwest UX (User Experience) conference where I took a workshop on improvisation from Mike Gorgone, a computer scientist, who learned improvisation from both The Second City and iO improv companies. His focus has been applying improvisation techniques to help UX designers produce digital products. Interestingly, Dick Costolo, the former CEO of Twitter, was a computer scientist and an accomplished improviser who studied improv while living in Chicago as well.

In February-March, 2019, I worked at Miras University in Kazakhstan on a project entitled “Multilingual Education via Interactive & Cognitive Teaching”, which was out of their Computer Science and Information Technology department. I conducted improv workshops for both their technical faculty, English faculty, and also elementary and high school English teachers from the region.

## 2. BASIC RULES OF IMPROVISATION

The seven rules used here come from The Second City and are as follows: listen, use “Yes, and”, create an ensemble, co-create the story, be willing to change, use failure to improve, and follow the follower. Improvisation is fluid, which makes it a good tool for creative problem-solving.

Improvisation can be thought of as a method where the group begins in chaos, which gives rise to order (and, eventually, return to chaos).

Active listening is a “muscle” that needs to be exercised and requires the following focus areas: the words themselves, the emotion behind the words, the perceived intent of the words, and that the listener is in the moment and not thinking of what s/he may be planning to say. Listen and only respond to what is heard. When someone speaks to you, treat it as a special gift and acknowledge it as being so.

By saying “yes”, one affirms the speaker and by saying “and” (as opposed to “but”), one indicates the intention to build on the speaker’s words. In this way, a story is built brick by brick. Such a method can take the performers (or problem solvers) to unexpected places and it may make one feel a bit at ease; however, it is a way of keeping a clear line of communication open. Affirm, build, and be non-judgmental are keys to this phase of improvisation. In the early days of IBM, the company’s motto was “THINK”, but in improvisation it is best to “don’t bother to think” as some of the best lines and ideas often come when performers say what is on their minds at that moment.

The word “ensemble” is purposefully chosen as opposed to the word “team”. An ensemble of musicians, for instance, is brought together with the idea that it will create a beautiful musical piece: it is the whole that is important, not the individual. The word team is often used to mean this as well; however, in terms of athletics, teams imply that some level of competition was needed in order to be a member of such teams, and that one can easily be replaced if injured. An ensemble preserves the talent and the collaborations are highly synergistic. In order to be successful, ensembles must be in the moment, give and take, and surrender the right to be right (the last one is hard even in brainstorming sessions).

The focus here needs to be on finding “the idea”, not one’s own idea. This means that one must be willing to cede control, which can be difficult. During this time, one needs to avoid asking questions as this can be seen as: blocking the idea from being developed, an attempt to force justification, avoidance, or just not wanting to participate. During this time of co-creation, a few basic rules of civility and behavior are needed. Basically, share the stage and treat every member with respect. Finally, it is important to remember that while some words or sayings may be funny to one person, they can be interpreted by another as having a negative connotation.

Changing oneself is very difficult and changing others or an organization is nearly impossible without first focusing on oneself. It also involves risk taking on one’s part. Improv can help in developing this as one is constantly faced with new ideas that s/he must adopt in order to keep a story moving forward. Taking risks has become a normal part of Silicon Valley and the startup mentality as a whole for quite some time. While failure is possible, the potential of a big monetary payoff seems to make it worthwhile. In the improv world, risk-taking and

failure are how actors improve their future performances. Improv also helps one learn how to fail in public, which can be very embarrassing and even emotionally upsetting if one is not trained for the possibility.

It is not about being the best or perfect at that moment, as much as it is about being in the moment and embracing the concept of life-long learning. In the classes attended by the author, we used the word “Ahougá” when we made a mistake or faux pas of some sort. There are some improvisation exercises which have the purpose to force mistakes on the team members. This teaches one how to move forward when a mistake is made without getting caught up in judgmental back-thinking.

Spolin stated “Don’t initiate! Follow the initiator! Follow the follower.” All of us have the ability to lead, and there are times when it is better for executives to step back and allow others to take over. Let things grow organically and see where they take the ensemble. Too much direction from the top tends to negate ideas from the bottom and can stifle innovation.

### 3. LITERATURE REVIEW

There is a fair amount of relevant research, articles, and talks on the use of improvisation outside of the theater. The articles reviewed show that improvisation can improve one’s communication, collaboration, creativity, and problem-solving skills.

1. “Beyond Productivity: Information, Technology, Innovation, and Creativity” edited by William J. Mitchell, Alan S. Inouye, and Marjory S. Blumenthal [10]. This is a rather lengthy book with a theme revolving around the importance of combining both the arts and STEM disciplines in order for the next generation of problem solvers and innovators to be successful. It is stated that this is “also an opportunity for each field to gain fresh, sometime uncomfortable, perspectives on itself.” It is pointed out that Pixar co-founder, Ed Catmull, embraced improvisation as a way to teach better collaboration within teams. The authors state that improvisation has a place in human-computer interaction or HCI. Finally, there are several references made to George E. Lewis, a professor of American Music at Columbia University, who said, “Improvisation is about finding structure, not imposing it.”
2. “Changes in Anxiety Following a Randomized Control Trial of a Theatre-based Intervention for Youth with Autism Spectrum Disorder” by Blythe A. Corbett, PhD, Scott D. Blain, Sara Ioannou, and Maddie Balser [3]. Their work indicates that improvisation techniques can be used to help people on the autism spectrum disorder (ASD). They point out a number of studies and suggest that additional work in this area is needed. They created the Social Emotional Neuroscience Endocrinology (SENSE) lab “to better understand the social and emotional functioning of children

- with autism and related neurodevelopmental disorders and to translate findings into meaningful approaches and interventions.”
3. “Education in Professionalism: Improvisation” by Richard B. Gunderman, MD, PhD [6]. This paper looks at how Dr. David Fessel, MD, a musculoskeletal radiologist and faculty member at the University of Michigan, took classes on improvisation and eventually became a member of The Second City Conservatory. His primary reason was to overcome his initial fear of giving presentations, but eventually he realized that there was much more to this. He learned that improvisation reaches far beyond that of comedy and can be applied to the medical profession in terms of teamwork, collaborative communication, relationship building, learning, and self-awareness.
  4. “Humour-in-the-loop: Improvised Theatre with Interactive Machine Learning Systems” by Kory Mathewson [9]. Mathewson is an accomplished improviser and, in 2015, he began his work on creating an AI system that could engage in improvisation with a human. He developed a set of Turing test rules that could be used to show success. He collaborated with Rapid Fire Theatre on this project and it is also here where he created the improvised theater experiment called Improbotics.
  5. “Improv for Effective Collaboration Innovation?” by Jeannie Kristufek [4]. In this paper, the author presents several exercises used at IBM that she adapted from the improv world to help technical team members work more creatively and efficiently as part of a collective whole. A key component of improvisation is that it is a team effort and that no “star” exists. Everyone is tasked with contributing to the task at hand and one must listen to the other team members even if there is disagreement. Using a “Yes, and” approach allows team members to avoid the value judgment of ideas early on and keeps the discussion going. Several other key components of improv that she points out are: “show, don’t tell”, take risks, be in the moment and the know.
  6. “Improvisational Computational Storytelling in Open Worlds” by Lara J Martin, Brent Harrison, and Mark Riedl [8]. Improvised storytelling allows the actors to place their characters in whatever universe and situation they wish based upon a simple suggestion from someone in the audience: in essence, what is referred to as an open world. The authors are researching how to do the same with an AI system that would be capable of improvising with humans in real-time alongside humans. They offer two possible approaches to help solve this problem: plot graph and neural network. Neither solution is perfect and much more needs to be done, but the goal is to develop an AI system that can communicate with humans at even some of the most abstract worlds that improvisers are good at creating.
  7. “Improvisation: Methods and Models” by Jeff Pressing [11]. This paper was written in 1987 and, although it is mostly on musically improvisation, many of the ideas

could easily be applied to theater improvisation as well. Even more interesting is the section on artificial intelligence that Pressing examines. Pressing notes: “There is traditionally no explicit mention of improvisation in the field. In making such a link, it seems clear that the successful application of AI concepts to improvisation rests to a large degree on the appropriateness of considering improvisation to be a kind of problem-solving.” He proceeds to discuss how improvisation is really a form of problem-solving in terms of a chosen search method and selection of an acceptable solution through problem reduction.

8. “The Use of Improvisational Theater Training to Reduce Social Anxiety in Adolescents” by Peter Felsman, Colleen M. Seiferta, and Joseph A. Himle [4]. This was a study conducted in Detroit in cooperation with The Improv Project, which is a part of the Detroit Creativity Project (DCP). The authors state, “This study is the first to test the efficacy of a school-based improvisational theater program as a mental health intervention, and it offers positive results.” The focus was on those youth who suffer from a recognized social anxiety disorder such as depression or social avoidance. I met with Beth Hagenlocker who is a co-founder of the DCP and in charge of the day-to-day operation of The Improv Project. During a 10-week period, students transition from the basic improv exercises to a graduation performance. There are others doing similar things such as Urban Improv in Boston, which is a part of Rehearsal for Life.
9. “Whose Classroom Is It, Anyway? Improvisation as a Teaching Tool” by Ronald A. Berk and Rosalind H. Trieber [1]. The focus was on using improvisation as a better way to engage and teach the Net Generation students. The authors used the following four improv exercises to demonstrate how this can be done: One Word at a Time/One Sentence at a Time, Speech Tag, Freeze Tag, and Gibberish Expert Interview. Their four reasons for using improvisation in the classroom are: 1. it is consistent with the characteristics of the current generation of students; 2. it taps into students’ multiple and emotional intelligences; 3. it fosters collaborative learning; and 4. it promotes deep learning through the active engagement.

#### 4. IMPROV EXERCISES

Below is a look at the exercises used by the author along with a brief summary.

Table 1  
*Improv Exercises Used by the Author*

	Most exercises can be found in the Improv Encyclopedia, Improv Wiki, Drama Toolkit, and Learnimprov.com although the names may be different
1	Shake 8s; Circle up -- A great warm-up exercise.
2	Zip-zap-zoop; Woosh-bang-pow -- Any three-word/sound combination is fine.
3	These are five things -- "Don't bother to think"; just do.
4	Follow the follower or the leader -- Great exercise and fun as people imitate others.
5	Last word spoken = first word spoken -- Forces one to concentrate.
6	New choice, redo -- I refer to this game as "refactor" which is a term used by programmers when they rewrite some of their code.
7	Double-link list -- In a circle, point to someone saying a "noun" you think of. Once the loop has gone around, repeat the cycle using the same words. Build on this by adding another round with a different set of words to see how well you can remember your word(s) or that/those of others. Use "pointers" (fingers) for navigation of the circuit.
8	Zombie/hacker attack (10-14 people) -- Have the same number of chairs as participants, which means one will be empty. The zombie walks slowly toward the empty chair while the others create a team strategy to cover.
10	Botnet, distributed denial of service or DDoS attack -- A winker (bot herder) creates a botnet that s/he can later command to attack a victim.
11	Bomber-protector -- Use 8-14 people; each person identifies one person as his/her bomber (therefore to be avoided) and another person as his/her protector (therefore wanting to stay as close as possible to).
12	Two- or three-headed experts -- Answer a question with either one word or one sentence at a time.
13	Four squares -- Four players with the front two being given a topic they will discuss; the entire square rotates to allow all actors the opportunity.
14	Circle story -- Any topic you want and everyone contributes to the story.
15	Emotion zone -- Allows for team members to work with different emotions.
16	Status game -- Use a deck of cards. Treat others based upon their perceived status.
17	Alphabet circle -- Use the alphabet to tell a story.
18	Knife-cat-baby throwing -- This is about object work and concentration.

## 5. CONCLUSIONS

There are anecdotal works, testimonials, and research that support how improvisation can help everyone become better researchers, technologists, teachers, entrepreneurs, team members, and humans. Since 2016, Northeastern University's computer science majors have been required to take a drama class which includes improvisation. [2] This is done, in part, in

an attempt to “robot-proof” their majors, as well as giving them better teamwork and creative problem-solving skills. The area of AI research has begun to look at how improvisation theater can be used to help develop better AI systems.

## REFERENCES

- Berk, R. A., & Trieber, R. H., (2009). *Whose Classroom Is It, Anyway? Improvisation as a Teaching Tool*, *Journal on Excellence in College Teaching*, v20 n3, pp. 29-60.
- Castellanos, S. (May 14, 2019). Oh, My God, Where Is This Going?’, When Computer-Science Majors Take Improv, *The Wall Street Journal*.
- Cobbett, BA et al. (Apr 12, 2017). Changes in Anxiety Following a Randomized Control Trial of a Theatre-based Intervention for Youth with Austim Spectrum Disorder, *PubMed*.
- Felsman, et al. (2018.12.001). The Use of Improvisational Theater Training to Reduce Social Anxiety in Adolescents. *The Arts in Psychotherapy*. <https://doi.org/10.1016/j.aip>
- Gunderman, R. (2016). *Education in Professionalism: Improvisation*. The Association of University Radiologists, Elsevier Inc.
- Kristufek, J. (2008). *Improv for Effective Collaborative Innovation?* WEPAN.
- Kuritz, P. (1988). *The Making of Theatre History*, Prentice Hall.
- Martin, L. J., Harrison, B., Riedl, M. (2016). *Improvisational Computational Storytelling in Open Worlds*, Lecture Notes in Computer Science book series, International Conference on Interactive Digital Storytelling (ICIDS).
- Mathewson, K. (2019). *Humour-in-the-loop: Improvised Theatre with Interactive Machine Learning Systems*, PhD thesis. Re [korymathewson.com/category/academic/computer-science/](http://korymathewson.com/category/academic/computer-science/)
- Mitchell, W, Inouye, A, and Blumenthal, M. (editors) (2003). *Beyond Productivity: Information, Technology, Innovation, and Creativity* (ITCP), National Academies Press.
- Pressing, J. (1987). *Improvisation: Methods and Models, Generative Processes in Music*, ed. J Sloboda, Oxford University Press.
- Sanders, B. (1995). *Sudden Glory: Laughter as Subversive History*, Beacon Press.



# Gamificación en la educación superior

Elvira Rincón-Flores

elvira.rincon@tec.mx/ Tecnológico de Monterrey, México

Recepción: 1-8-2019 / Aceptación: 21-8-2019

**RESUMEN.** La gamificación es una estrategia que surgió en la milicia y que fue aplicada posteriormente en el área comercial con el propósito de fortalecer la fidelidad de los consumidores. En los últimos años ha sido adoptada por los educadores de diversas partes del mundo con el propósito de enganchar y motivar a los estudiantes en su proceso de aprendizaje, tanto en la educación a distancia como en los cursos presenciales. La gamificación toma elementos del juego en contextos que no son del juego, por ello resulta un recurso valioso que permite integrar aspectos cognitivos, sociales y emotivos, los cuales favorecen el aprendizaje. El propósito del presente trabajo es mostrar cómo estos elementos son percibidos por los estudiantes y cómo son capitalizados en el proceso de enseñanza-aprendizaje cuando se aplican actividades que incluyen gamificación basadas en retos.

**PALABRAS CLAVE:** gamificación basada en retos, innovación educativa

## Gamification in Higher Education

**ABSTRACT.** Gamification is a strategy that emerged from armed forces and was developed later in business for strengthening consumer's loyalty. In recent years, it has been adopted worldwide by educators to engage and motivate students during their learning process through both distance education and face-to-face courses. Gamification uses gaming elements in non-gaming contexts. Therefore, it is a valuable resource that allows the integration of cognitive, social and emotional aspects which favor learning. The purpose of this paper is to show how these elements are perceived by students and how they are capitalized on the teaching-learning process when activities that include challenge-based gamification are carried out.

**KEYWORDS:** challenge-based gamification, educational innovation

## 1. INTRODUCCIÓN

### 1.1 Innovación educativa para mejorar el proceso enseñanza-aprendizaje

La innovación educativa ha tomado fuerza en las últimas décadas y a la vez ha dado lugar a opiniones encontradas (Fidalgo-Blanco, 2014). Pues no se trata de cambiar por cambiar, es decir, introducir una tecnología novedosa o alguna estrategia didáctica al proceso de enseñanza-aprendizaje no siempre resultará en una innovación educativa (Fidalgo-Blanco y Sein-Echaluce, 2014; García-Peñalvo, 2015). Para que dicha innovación se dé, los cambios deben estar sustentados en la enseñanza, que el propósito sea mejorar el aprendizaje y que, además, esa innovación sea evaluada con el propósito de verificar que se mejoró el proceso de enseñanza-aprendizaje (García-Peñalvo, García de Figuerola y Merlo, 2010; Rincón-Flores, Gallardo y de la Fuente, 2018).

Con el propósito de evaluar las innovaciones en la educación superior, Borrego, Froyd y Hall (2010) establecieron cuatro criterios: a) existen investigaciones en torno a la innovación que se adopta; b) las innovaciones han sido adoptadas universalmente en la institución; c) existe evidencia de su efecto en el aprendizaje o retención de alumnos; y d) la innovación se distingue de otras innovaciones. Por su parte, Zabalza Beraza y Zabalza Cerdeiriña (2012) sostienen que existen tres elementos esenciales en el desarrollo de una innovación: la apertura, la actualización y la mejora de la calidad. La apertura es la capacidad de adaptación, así como el desarrollo o mejora de actitudes, conocimientos, destrezas y recursos. La actualización consiste en estar al corriente de los últimos avances relacionados a la innovación implementada, mientras que la mejora de la calidad consiste en evaluar en qué medida la innovación mejora el proceso. Estos criterios son interesantes y pueden ser útiles para evaluar diversas innovaciones educativas de cualquier área del conocimiento.

En la enseñanza de las matemáticas se han hecho esfuerzos para incluir estrategias didácticas innovadoras con el objetivo de mejorar el aprendizaje de los alumnos y, más aún, las actitudes hacia esta disciplina (Gómez Chacón, 2000; Rincón-Flores, Cienfuegos y Carrillo, 2015; Tobias y Weissbrod, 1980). Algunas estrategias didácticas que se han aplicado de manera exitosa en la enseñanza de las matemáticas a nivel superior son: aprendizaje activo a través de la técnica de la pregunta (Rincón-Flores, Cienfuegos, Galván y Fabela, 2014), modelación matemática (Blum y Borromeo, 2009; Rincón-Flores, Illanes y Gallardo, 2015), aprendizaje lúdico (Kebritchi, Hirumi y Bai, 2010; Muñoz-Rodríguez, Alonso y Rodríguez-Muñoz, 2014; Chamoso, Durán, García, Martín y Rodríguez, 2004) y gamificación (Rojas-López, Rincón-Flores, Mena, García-Peñalvo y Ramírez-Montoya, 2019; Caponetto, Earp y Ott, 2014; Villalustre y del Moral, 2015; Hamari *et al.*, 2016; Rincón-Flores e Illanes, 2015), entre otros.

Con respecto a la innovación, Brousseau (1990), reconocido investigador francés en matemática educativa, sostiene que es un mecanismo didáctico que propone estrategias que

funcionan y que son fácilmente comunicables a los demás, pues innovar no solo implica mejorar las técnicas de enseñanza sino generar cambios auténticos que favorezcan el aprendizaje. En el caso de la enseñanza de las matemáticas es de suma importancia implementar un cambio en los procesos formativos que mejoren las actitudes de los estudiantes hacia las matemáticas y que con ello se produzcan condiciones para un aprendizaje más significativo.

## 1.2 Gamificación en la enseñanza de las matemáticas en la educación superior

En los últimos años la estrategia de gamificación ha ido tomando más fuerza gracias al elemento lúdico que la caracteriza, pues el juego es una actividad universal presente en la historia de la humanidad en diversas áreas y disciplinas (Chamoso *et al.*, 2004). Huizinga (1938), filósofo e historiador holandés, quien fue el primero en estudiar el fenómeno de la lúdica en la vida del hombre en los años cincuenta, sostuvo que el juego es más antiguo que la cultura misma, pues el ser humano, al igual que los animales, ha jugado desde siempre y se percibe como una actividad útil que no sólo ofrece diversión, relajación o entretenimiento, sino aprendizajes valiosos.

A lo largo de la historia, las matemáticas han tenido un componente lúdico. Por ejemplo, Pascal, Fermat y Gauss participaban en duelos matemáticos, lo que podría traducirse como una insipiente gamificación basada en retos. Chamoso *et al.* (2004) establecieron un conjunto de características fundamentales del juego de las cuales se destacan: el carácter lúdico, las reglas propias y el carácter competitivo. La gamificación en la educación surge como un derivado del aprendizaje lúdico. La definición más generalizada de gamificación es que usa los elementos del juego en contextos que no son de juego (Deterding, Khaled, Nacke y Dixon, 2011). El término fue acuñado por Nick Pelling en el 2002, aunque tomó popularidad hasta el 2010, gracias a los sistemas de recompensa en entornos digitales que ofrece el sector empresarial y comercial a sus consumidores (Rodríguez y Santiago, 2015).

Años más tarde, la gamificación comenzó a incorporarse en el entorno educativo alrededor del mundo, lo que ha generado mayor interés a las comunidades académicas. En la educación se utiliza con el propósito de colocar al estudiante en escenarios que impliquen el desarrollo de retos y misiones atractivas que aumenten su nivel de compromiso y competitividad (Hanus y Fox, 2015), ya sea como recurso digital diseñado para un propósito didáctico o bien para tomar sus elementos e integrarlos en actividades educativas a través de simulaciones (Marín Díaz, 2015). En este sentido, el aprendizaje basado en retos puede ser un elemento que complemente a la gamificación y viceversa. Un reto es un desafío que los alumnos resuelven a través de la aplicación de lo que han aprendido (Hamari *et al.*, 2016). Así, ambas estrategias pueden converger en lo que se ha denominado gamificación basada en retos, pues la gamificación motivaría el compromiso del estudiante (Rojas-López y Rincón-Flores, 2018). El aprendizaje basado en retos motiva su creatividad en la resolución de problemas determinando el nivel de aprendizaje alcanzado (Martin, Rivale y Diller, 2007).

Diversas publicaciones muestran los beneficios de la gamificación en la educación. Por ejemplo, Kebritchi *et al.* (2010) se dieron a la tarea de buscar estudios empíricos en donde se utilizaba la gamificación como estrategia de aprendizaje de las matemáticas, encontrando al menos dieciséis investigaciones, del 2003 al 2007, de las cuales once tuvieron resultados positivos y cinco tuvieron tanto resultados positivos como negativos. Para Hanus y Fox (2015), la gamificación es un importante aliado ya que puede ser aplicado en la educación como medio para motivar y utilizar nuevos caminos para disfrutar de actividades que suelen ser tediosas, lo que resulta conveniente cuando se abordan contenidos matemáticos complejos.

La gamificación posee elementos de gran valor en la formación integral de los estudiantes. Domínguez *et al.* (2013) y Nisbet y Williams (2009) coinciden en que el juego constituye una oportunidad única para integrar los aspectos cognitivos, afectivos y sociales. Sin duda estos aspectos deben ser considerados en el diseño de toda actividad gamificada. El aspecto cognitivo se da cuando el alumno obtiene retroalimentación inmediata y se le dan varios intentos de tal manera que se le conduce a un proceso metacognitivo o bien cuando se le enfrenta a un reto (Rojas-López, Rincón-Flores, Mena, García-Peñalvo y Ramírez-Montoya, 2019). El aspecto emotivo se da cuando el estudiante obtiene un reconocimiento por su logro (Mekler, Brühlmann, Tuch y Opwis, 2017) y el aspecto social sucede cuando los logros son socializados a través de un tablero de liderazgo o bien cuando los estudiantes trabajan colaborativamente para lograr un reto o misión (Domínguez *et al.*, 2013; Hanus y Fox, 2015).

El propósito del presente trabajo es mostrar la experiencia didáctica de gamificación basada en retos en un curso de cálculo en el nivel superior, así como sus ventajas desde la perspectiva del estudiante y del profesor, a partir de las tres dimensiones que toda experiencia gamificada debe integrar: cognitiva, social y emotiva.

## 2. METODOLOGÍA

Es una investigación cuantitativa y exploratoria, en la que se aplicó un cuestionario con escala de Likert para conocer la relación de la gamificación y las tres dimensiones: cognitiva, social y emotiva. Se trabajó con dos grupos, cada uno con treinta estudiantes, del curso Matemáticas II para Ingeniería.

La investigación se llevó a cabo a lo largo del semestre, de enero a mayo del 2018, en cuatro etapas principales:

**Primera etapa:** los estudiantes conformados en binas seleccionan un avatar y es pegado en un tablero de liderazgo. El propósito del uso del avatar es que los estudiantes puedan estar en anonimato y que no se sientan expuestos o ansiosos al ver los avances de los otros equipos (véase la figura 1).



Figura 1. Tablero de liderazgo cuando la actividad gamificada basada en retos concluyó

*Segunda etapa:* previamente se prepararon tres retos con nivel de dificultad progresivo. El equipo no puede pasar al siguiente nivel si no ha terminado el anterior. En esta etapa a los estudiantes se les proporciona retroalimentación inmediata, si no ha respondido correctamente puede volver a intentar, lo interesante es que en este momento se detona un proceso metacognitivo en el que el estudiante reflexiona sobre sus aciertos y errores.

En la figura 2 se muestra el primer reto de la sesión gamificada basada en retos, el cual consistió en que encontrarán la integral general para obtener el volumen de un sólido de revolución que gira en torno al eje de las  $x$ , utilizando la estrategia de la toma de un diferencial (Alanís y Salinas, 2010; Rincón-Flores e Illanes, 2015) asociándolo al cálculo geométrico del volumen de un cilindro. El reto consistió en obtener el volumen de una copa.

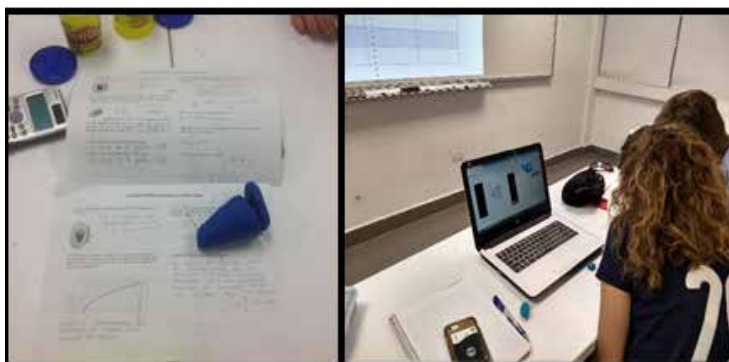


Figura 2. Ejemplo del nivel 1 de la actividad gamificada basada en retos

*Tercera etapa:* una vez que la actividad termina, en la misma clase, si el tiempo lo permite, o en la siguiente, se discuten todos los retos para reforzar la actividad.

*Cuarta etapa:* se aplicó el cuestionario sobre dimensiones de la gamificación.

### 3. RESULTADOS

Se presentan los resultados cuantitativos del instrumento con escala de Likert. El cuestionario fue respondido por 52 estudiantes de manera voluntaria, 26 de cada grupo. El cuestionario estuvo conformado por doce preguntas, cuatro asignadas a cada dimensión. A continuación, se muestran los resultados.

La dimensión cognitiva trata de evaluar cómo percibe el estudiante la relación de la gamificación con respecto a los procesos cognitivos. Se presentan los resultados de dos preguntas (véase la figura 3).

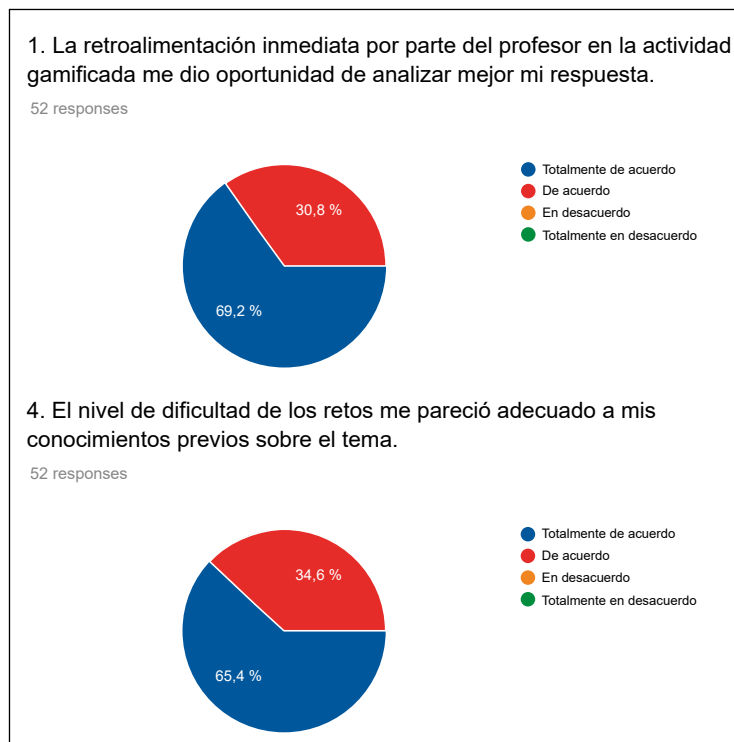


Figura 3. Resultados parciales de la dimensión cognitiva (resultados recabados por el autor)

Los resultados muestran que los estudiantes percibieron que la gamificación basada en retos favoreció su aprendizaje, ya que al recibir una retroalimentación inmediata y al tener oportunidad de volver a intentar el reto, les dio la oportunidad de repensar y reflexionar sobre su respuesta, dando lugar a un proceso metacognitivo, esto coincide con lo encontrado por Rojas-López *et al.* (2019). Por otro lado, también percibieron que el reto era adecuado a su nivel de conocimientos, lo cual es un factor importante ya que si fuese elevado provocaría que la motivación disminuyera, caso contrario si fuese fácil, se perdería la sensación de logro.

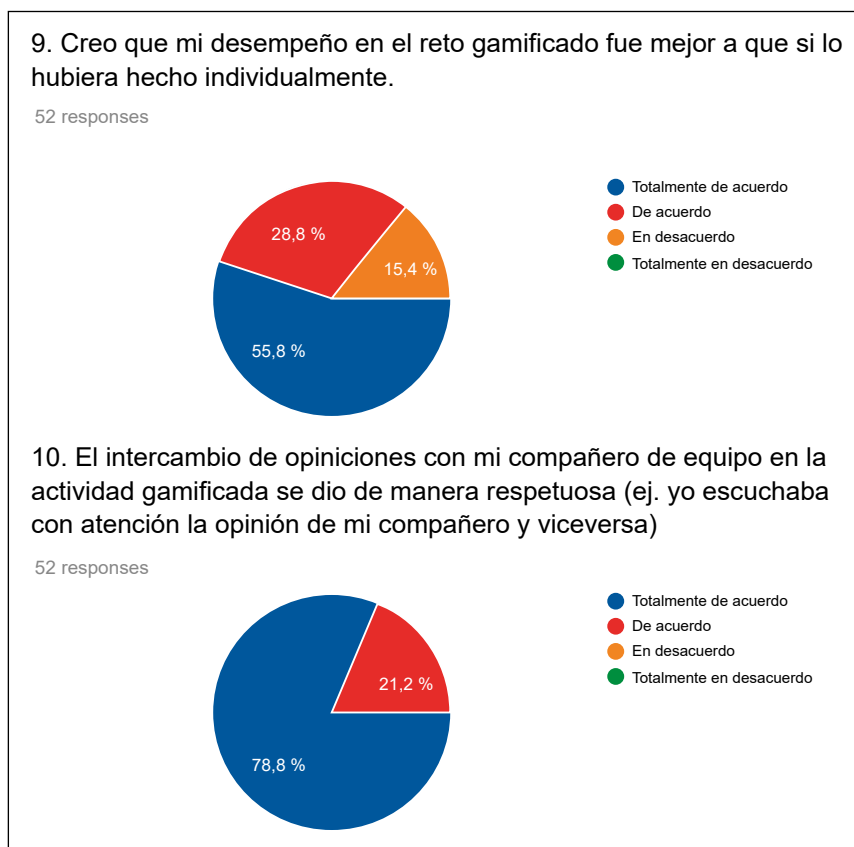


Figura 4. Resultados parciales de la dimensión social (resultados recabados por el autor)

En cuanto a la dimensión social, se puede ver que este componente le da un matiz importante a la actividad gamificada basada en retos, pues los estudiantes percibieron que su desempeño fue mejor gracias al intercambio de ideas con sus compañeros de equipo (véase la figura 4). Al mismo tiempo, al llevarse a cabo la actividad en un contexto de respeto también promueve valores y actitudes, los cuales también son importantes en la formación del futuro

profesional. Esto coincide con lo señalado por Caponetto *et al.* (2014), Domínguez *et al.* (2013), y Lee y Hammer (2011).



Figura 5. Resultados parciales de la dimensión emotiva (resultados recabados por el autor)

La tercera dimensión, la emotiva, se refiere a las emociones que toda actividad relacionada al juego puede provocar en las personas, lo cual también constituye un elemento importante y es inherente a las actividades gamificadas, incluso neurocientíficos como Bueno i Torrens (2019) y Mora (2014) sostienen que las emociones favorecen el aprendizaje. Los resultados de la figura 5 muestran que los estudiantes se percibían contentos cada vez que su equipo terminaba un reto y emocionados cuando veían su avance en el tablero de liderazgo. El componente emocional tiene una influencia positiva en los aspectos motivacionales (Rojas-López *et al.*, 2019) y contribuye a que los estudiantes estén en un estado de *flow* (Hamari *et al.*, 2016).

Finalmente, a partir de los resultados obtenidos, se puede afirmar que la gamificación es una estrategia didáctica innovadora porque cumple con los criterios señalados por Borrego *et al.* (2010); Zabalza Beraza y Zabalza Cerdeiriña (2012).



#### 4. CONCLUSIONES

La gamificación en la educación puede ser una herramienta didáctica valiosa que permite conjugar diversos aspectos que son importantes en la formación del estudiante, lo cognitivo, lo social y lo emotivo. Lo cognitivo porque tiene que ver directamente con su aprendizaje, más aún, con un aprendizaje significativo; lo social porque representa una oportunidad para la discusión e intercambio de aprendizaje, al tiempo que se promueven actitudes y valores; lo emotivo porque este componente influye en la motivación, es decir, mantiene la atención del estudiante en un plazo mayor. No obstante, es importante que el diseño de las actividades gamificadas basadas en retos sean creativas y de interés para los estudiantes, además de que sean alcanzables por ellos, pues un diseño equivocado puede afectar significativamente el propósito de la actividad.

#### REFERENCIAS

- Alanís, J. A., y Salinas, P. (2010). Cálculo de una variable: Acercamientos newtoniano y leibniziano integrados didácticamente. *El Cálculo y su Enseñanza*, 2, 1-14. Recuperado de [http://mattec.matedu.cinvestav.mx/el\\_calculo/index.php?vol=2&index\\_web=8&index\\_mgzne](http://mattec.matedu.cinvestav.mx/el_calculo/index.php?vol=2&index_web=8&index_mgzne)
- Blum, W. y Borromeo Ferri, R. (2009). Mathematical modelling: Can it be taught and learnt? *Journal of Mathematical Modelling and Application*, 1(1), 45-58.
- Borrego, M., Froyd, J. E., y Hall, T. S. (2010). Diffusion of engineering education innovations: A survey of awareness and adoption rates in U.S. engineering departments. *Journal of Engineering Education*, 9(3), 185-207.
- Brousseau, G. (1990). Utilidad e interés de la didáctica para un profesor (2.ª parte), *SUMA*, (5), 5-12.
- Bueno i Torrens, D. (2019). *Neurociencia para educadores*. Barcelona: Octaedro.
- Caponetto, I., Earp, J., y Ott, M. (2014). Gamification and education: A literature review. *8th European Conference on Games Based Learning*, 1, 50-57. Recuperado de [https://pdfs.semanticscholar.org/2b9b/64350c1d2d6f9103b8505612e98afb1da3bb.pdf?\\_ga=2.173855587.1792265263.1579307919-635103138.1579127417](https://pdfs.semanticscholar.org/2b9b/64350c1d2d6f9103b8505612e98afb1da3bb.pdf?_ga=2.173855587.1792265263.1579307919-635103138.1579127417)
- Chamoso, J., Durán, J., García, J., Martín, J., y Rodríguez, M. (2004). Análisis y experimentación de juegos como instrumentos para enseñar matemáticas. *SUMA*, (47), 47-58.
- Deterding, S., Khaled, R., Nacke, L. E., y Dixon, D. (2011). Gamification: toward a definition. *Chi 2011*, 12-15. doi:978-1-4503-0268-5/11/0

- Domínguez, A., Saenz-de-Navarrete, J., de-Marcos, L., Fernández-Sanz, L., Pagés, C., y Martínez-Herráiz, J. J. (2013). Gamifying learning experiences: Practical implications and outcomes. *Computers and Education*, 63, 380-392. doi:10.1016/j.compedu.2012.12.020
- Fidalgo-Blanco, Á. (2014). Innovación educativa en la sociedad del conocimiento. *Teoría de la Educación y Sociedad de la Información TESI*, 15(3), 1-4.
- Fidalgo-Blanco, Á., y Sein-Echaluce, M. L. (2014). Educational innovation. *Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'14)*, 65-67. doi:10.1145/2669711.2669880
- García-Peñalvo, F.J., García de Figuerola, C., y Merlo, J.A. (2010). Open knowledge: challenges and facts. *Online Information Review*, 34(4), 520-539. doi:10.1108/14684521011072963
- García-Peñalvo, F. J. (2015). Mapa de tendencias en innovación educativa. *Education in the Knowledge Society*, 16(4), 6-23. doi:10.14201/eks2015164623
- Gómez Chacón, I. M. (2000). *Matemática emocional: los afectos en el aprendizaje matemático*. Madrid: Narcea.
- Gros Salvat, B., y Lara Navarra, P. (2009). Estrategias de innovación en la educación superior: el caso de la Universitat Oberta de Catalunya. *Revista Iberoamericana de Educación*, (49), 223-245. Recuperado de <https://rieoei.org/historico/documentos/rie49a09.pdf>
- Hamari, J., Shernoff, D. J., Rowe, E., Coller, B., Asbell-Clarke, J., y Edwards, T. (2016). Challenging games help students learn: an empirical study on engagement, flow and immersion in game-based learning. *Computers in Human Behavior*, 54, 170-179. doi:10.1016/j.chb.2015.07.045
- Hanus, M., y Fox, J. (2015). Assessing the effects of gamification in the classroom: a longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance, *Computers & Education*, 80, 152-161.
- Huizinga, J. (1938). *Homo Ludens*. Países Bajos: Penguin Random House Grupo Editorial.
- Ke, F., y Grabowski, B. (2007). Gameplaying for maths learning: cooperative or not? *British Journal of Educational Technology*, 38(2), 249-259. doi:10.1111/j.1467-8535.2006.00593.x
- Kebritchi, M., Hirumi, A., y Bai, H. (2010). The effects of modern mathematics computer games on mathematics achievement and class motivation. *Computers & Education*, 55(2), 427-443.
- Lee, J., y Hammer, J. (2011). Gamification in education: what, how, why bother? *Academic Exchange Quarterly*, 15(2), 1-5.

- Marín Díaz, V. (2015). La gamificación educativa. Una alternativa para la enseñanza creativa. *Digital Education Review*, (27), 5-8. Recuperado de [http://revistes.ub.edu/index.php/der/article/view/12486/pdf\\_1](http://revistes.ub.edu/index.php/der/article/view/12486/pdf_1)
- Martin, T., Rivale, S. D., y Diller, K. R. (2007). Comparison of student learning in challenge-based and traditional instruction in biomedical engineering. *Annals of Biomedical Engineering*, 35(8), 1312-1323. doi:10.1007/s10439-007-9297-7
- Mekler, E. D., Brühlmann, F., Tuch, A. N., y Opwis, K. (2017). Towards understanding the effects of individual gamification elements on intrinsic motivation and performance. *Computers in Human Behavior*, 71(C), 525-534. doi:10.1016/j.chb.2015.08.048
- Mora, F. (2014). *Neuroeducación*. Madrid: Alianza Editorial.
- Muñiz-Rodríguez, L., Alonso, P., y Rodríguez-Muñiz, L. J. (2014). El uso de los juegos como recurso didáctico para la enseñanza y el aprendizaje de las Matemáticas: estudio de una experiencia innovadora. *Revista Iberoamericana de Educación Matemática*, (39), 19-33. Recuperado de <http://www.fisem.org/www/union/revistas/2014/39/archivo6.pdf>
- Nisbet, S., y Williams, A. (2009). Improving students' attitudes to chance with games and activities. *Australian Mathematics Teacher*, 65(3), 25-37.
- Rincón-Flores, E. G., e Illanes, L. (2015). Aprendizaje gamificado en un curso de Cálculo para Ingeniería. *Memorias del II Congreso Internacional de Innovación Educativa*. Recuperado de <http://ciie.mx/memorias/>
- Rincón-Flores, E. G., Cienfuegos, D. E., y Carrillo, O. (2015). Aprendizaje activo y su efecto en las actitudes hacia las Matemáticas. *Memorias del II Congreso Internacional de Innovación Educativa*. Recuperado de <http://ciie.mx/memorias/>
- Rincón-Flores, E. G., Gallardo, K., y de la Fuente, J. M. (2018). Strengthening an educational innovation strategy: processes to improve gamification in calculus course through performance assessment and meta-evaluation. *International Electronic Journal of Mathematics Education*, 13(1), 1-11. doi:10.12973/iejme/2692
- Rincón-Flores, E. G., Cienfuegos, D. E., Galván, D., y Fabela, L. (2014). El aprendizaje activo como estrategia didáctica para la enseñanza del cálculo. *Acta Latinoamericana de Matemática Educativa*, (27), 499-506.
- Rodríguez, F., y Santiago, R. (2015). *Gamificación: cómo motivar a tu alumnado y mejorar el clima en el aula*. Barcelona: Grupo Océano.
- Rojas-López, A., Rincón-Flores, E. G., Mena, J., García-Peñalvo, F. J., y Ramírez-Montoya, M. S. (2019). Engagement in the course of Programming in Higher Education through the use of Gamification. *Universal Access in the Information Society (U AIS)*, 18, 583-597. doi:10.1007/s10209-019-00680-z

- Rojas-López, A., y Rincón-Flores, E. G. (2018). Gamification as Learning Scenario in Programming Course of Higher Education. Springer International Publishing AG, part of Springer Nature 2018, 1, 200-210.
- Tobias, S., y Weissbrod, C. (1980). Anxiety and mathematics: an update. *Harvard Educational Review*, 50(1), 63-70.
- Tsai, M.-J., Huang, L.-J., Hou, H.-T., Hsu, C.-Y., y Chiou, G.-L. (2016). Visual Behavior, Flow and Achievement in Game-Based Learning. *Computers & Education*, 98, 115-129. doi:10.1016/j.compedu.2016.03.011
- Villalustre Martínez, L., y del Moral Pérez, M. E. (2015). Gamificación: Estrategia para optimizar el proceso de aprendizaje y la adquisición de competencias en contextos universitarios. *Digital Education Review*, 27.
- Zabalza Beraza, M. A., y Zabalza Cerdeiriña, M. A. (2012). *Innovación y cambio en las instituciones educativas*. Rosario, Argentina: Homo Sapiens Ediciones.

# What Can Innovation in Engineering Education Do for You as a Student and What can You Do as a Student for Innovation in Engineering Education?

Carlos Alario-Hoyos  
calario@it.uc3m.es / Universidad Carlos III de Madrid. España

Recepción: 17-6-2019 / Aceptación: 3-7-2019

**ABSTRACT.** Innovation in education in general and innovation in engineering education in particular must be supported by properly collected and analyzed data to guide decision-making processes. Today it is possible to collect data from many more stakeholders (not just students), and also to collect much more data from each stakeholder. Nevertheless, low-level data collected by monitoring the interactions of the multiple stakeholders with learning platforms and other computing systems must be transformed into meaningful high-level indicators and visualizations that guide decision-making processes. The aim of this paper is to discuss some notable trends in data-driven innovation in engineering education, including 1) improvement of educational content; 2) improvement of learners' social interactions; 3) improvement of learners' self-regulated learning skills; and 4) prediction of learners' behavior. However, there are also significant risks associated with data collection and processing, such as privacy, transparency, biases, misinterpretations, etc., which must also be taken into account, and require creating specialized units and training the personnel in data management.

**KEYWORDS:** Data-driven innovation, learning analytics, digital education

¿Qué puede hacer la innovación de la educación de la ingeniería para el estudiante y qué puede hacer el estudiante para la misma?

**RESUMEN.** La innovación en la educación, en general, y la innovación en la educación de ingeniería, en particular, deben estar respaldadas por datos, debidamente recopilados y analizados para guiar los procesos de toma de decisiones. Hoy es posible recopilar datos de muchos grupos de interés (no solo estudiantes), y también recopilar muchos más datos de cada interesado. Sin embargo, los datos de bajo nivel recopilados al monitorear las interacciones de los múltiples interesados con las plataformas de aprendizaje y otros sistemas informáticos deben

transformarse en indicadores y visualizaciones de alto nivel que guíen los procesos de toma de decisiones. El objetivo de este documento es discutir algunas tendencias notables en la innovación basada en datos en la educación de ingeniería, que incluyen: 1) mejora del contenido educativo; 2) mejora de las interacciones sociales de los alumnos; 3) mejora de las habilidades de aprendizaje autorreguladas de los alumnos; y 4) predicción del comportamiento de los alumnos. Sin embargo, también existen riesgos significativos asociados con la recopilación y el procesamiento de datos, que incluyen privacidad, transparencia, sesgos, malas interpretaciones, etc., que también deben tenerse en cuenta y que requieren la creación de unidades especializadas y la capacitación del personal en la gestión de datos.

**PALABRAS CLAVE:** innovación basada en datos, análisis de aprendizaje, educación digital

## 1. INTRODUCTION

The rapid changes taking place in today's world have important consequences for education in general, and for engineering education in particular. Traditional “chalk and talk” teaching methodologies are not effective in engineering education, where it is necessary to combine the explanation of complex concepts with the practice of these concepts in realistic scenarios. In order to do this, it is necessary to apply learner-centered methodologies that promote active learning (Alario-Hoyos, Estévez-Ayres, Delgado Kloos, Villena-Román, Muñoz-Merino, et al., 2019), and that can be tailored to different education contexts, including face-to-face education in the classroom, online education, and blended (or hybrid) education (Pérez-Sanagustín, Hilliger, Alario-Hoyos, Delgado Kloos, & Rayyan, 2017).

Innovation in engineering education cannot take place without the support of a multidisciplinary team of specialized professionals. Instructional designers must advise teachers on how to redesign their traditional course to turn it into a blended learning experience, making the most of the available resources with the aim to promote active learning (Baepler, Walker, & Driessen, 2014). Pedagogues and psychologists must contribute with the understanding and development of self-regulated learning (SRL) skills to ensure student success, especially in engineering education where further development of SRL skills is required to succeed (Zimmerman, 2013). Developers must create applications and simulators to put into practice and evaluate the key concepts of each course. Researchers and data scientists must collect and analyze data, offering visualizations so that the main stakeholders of the educational process (students, teachers, managers, etc.) can make informed decisions. All in all, innovation in engineering education is a multidisciplinary effort in which all supporting professionals and actions must be aligned with the ultimate goal to benefit students.

Among the abovementioned professionals, the role of researchers and data scientists is becoming more and more important in the last few years. Over the course of time, many efforts to innovate in engineering education have been undertaken without considering the research advances in the field of technology-enhanced learning. Today, with the amount of data that can be collected for the particular context of each educational institution, there is no longer an excuse for not implementing data-driven decision-making processes at the different levels: 1) teachers must rely on data to improve their content and the methodologies used in their classes; 2) students must rely on data to detect their knowledge gaps and improve their SRL skills; and 3) managers must rely on data to detect unbalanced programs or unsatisfactory courses, and to offer personalized pathways for students.

The aim of this paper is to present and discuss some relevant trends in data-driven innovation in engineering education on the basis of recent research results. All the trends analyzed here have the student as the ultimate target stakeholder, so these trends are framed within the area known as learning analytics (Ferguson, 2012). However, the collection and processing of data for decision-making have a twofold side and are not without risks, some of which are

also discussed in this paper. This is precisely the structure followed in the rest of this paper, with section 2 discussing trends on data-driven innovation in engineering education, section 3 dedicated to discussing associated risks, and section 4 presenting the conclusions of this work.

## 2. SOME TRENDS ON DATA-DRIVEN INNOVATION IN ENGINEERING EDUCATION

One of the most important trends in data-driven innovation for engineering education is the improvement of educational content. Educational content may include, for example, video lectures, automatic correction exercises, and other additional resources (texts, animations, simulations, etc.). Nowadays, it is possible to collect low-level data related to the interaction of each learner with each educational resource, including (Ruipérez-Valiente, Muñoz-Merino, Leony, & Delgado Kloos, 2015) when the learner starts to play a video, when the learner stops a video, number of seconds of the video watched by the learner, when the learner attempts an automatic correction exercise, number of attempts in each automatic correction exercise by the learner, learning sequence followed by the learner when moving between educational resources, etc. With this low-level data, it is possible to detect videos or exercises that are not working correctly and that need to be improved. Some high-level indicators that contribute to detecting video lectures that need to be revised are: a high number of repetitions in a fragment of a video (which typically denotes a complex explanation or an error in that fragment), and a high percentage of students who do not watch a video from the beginning to the end (which typically denotes inappropriate content for the student's level or lack of engagement in the teacher's explanation). Some high-level indicators that contribute to detecting automatic correction exercises that need to be revised are: a very low number of incorrect answers in formative exercises (which typically denotes very simple exercises that may cause boredom and waste students' time), and a very high number of incorrect answers in summative exercises (which typically denotes very complex exercises that may cause frustration as students are not well prepared to solve those exercises).

Another important trend in data-driven innovation in engineering education refers to the improvement of social interactions among learners, and applies typically to courses with a very large number of students, either online or blended courses, and where teachers cannot provide personalized assistance due to the very large number of social interactions that take place in the course. Research in this line focused on characterizing the social interactions produced in a course, and on proposing methods and visualizations to help teachers make decisions about how to improve their course design. For example, there have been research studies which detected that the most appropriate tool to manage social interaction in courses with a very large number of students is the built-in forum provided by the learning platforms (Alario-Hoyos, Pérez-Sanagustín, Delgado-Kloos, Parada G., & Muñoz-Organero, 2014). Some other research studies focused on the identification of leaders within the community of learners, characterizing these leaders as the most active students in the course forum



(Alario-Hoyos, Muñoz-Merino, Pérez-Sanagustín, Delgado Kloos, & Parada G., 2016). This identification of leaders is important to facilitate teachers' work, as leaders can act as a bridge between the faculty and the rest of the students, even receiving special roles to be able to curate forum messages. Some other research studies focused on analyzing the overall class mood from social interactions, calculating the polarity of messages (positive, neutral, negative) posted by students in the course forum. The polarity of messages was calculated by applying word dictionaries and syntax rules, and the aim was to detect parts of the course in which the overall class mood was more positive or more negative to take corrective measures in the second case (Moreno-Marcos, Alario-Hoyos, Muñoz-Merino, Estévez-Ayres, & Delgado Kloos, 2018). Finally, all the data collected from social interactions can be used as input to develop chatbots or conversational agents programmed to give support to students in specific courses (Delgado Kloos, Catalán-Aguirre, Muñoz-Merino, Alario-Hoyos, 2018).

An additional relevant trend in data-driven innovation in engineering education is the study, characterization, and support of students' development of self-regulated learning (SRL) skills. SRL skills are particularly important in engineering education due to, among other things, the complexity of the contents and the permanent retraining demanded in today's engineers. Appropriate strategies to self-regulate each one's learning should be applied in every learning stage (before, during, and after each learning activity), including, for instance, setting reasonable and measurable objectives (before), seeking help when necessary (during), self-reflecting on the work done and the objectives achieved (after), etc. (Alonso-Mencía, Alario-Hoyos, Maldonado-Mahauad, Estévez-Ayres, Pérez-Sanagustín, & Delgado Kloos, 2019). The study and characterization of SRL skills in engineering courses led to the conclusion that strategies related to appropriate time management are the most problematic ones for learners. Therefore, specific interventions need to be done to facilitate time management, both when designing a course and when developing support tools (Alario-Hoyos, Estévez-Ayres, Pérez-Sanagustín, Delgado Kloos, & Fernández-Panadero, 2017). It is also important to support students' self-reflection by offering high-level visualizations based on low-level data with the aim to increase students' awareness on the desired level to be achieved and that of their classmates, both for an entire course and for each module or part of a course (Ruipérez-Valiente, Muñoz-Merino, Leony, & Delgado Kloos, 2015).

One last, but not least, relevant trend in data-driven innovation in engineering education is the prediction of students' behavior from previously collected data and appropriately trained machine learning models. The variables that are typically predicted through these models refer to students' partial or final grades in a course; and also to whether a student will abandon a course or not; and, by extension, to whether a student will abandon a complete study program (bachelor's degree or master's degree) or not. The aim of these prediction models is to take corrective measures in order to prevent students from failing a course, or from dropping out of a course or study program. Studies on prediction in education have detected that, in general, low-level data, such as the interaction with educational content (e.g., videos and exercises),

have greater predictive power than data collected from self-reported questionnaires (such as students' intentions and motivations) (Moreno-Marcos, Alario-Hoyos, Muñoz-Merino, & Delgado Kloos, 2018). There are still important gaps in this research line, including proposing generalizable models applicable to different educational contexts and areas of knowledge, and developing predictive models and tools for real-time data collection and processing in order to improve the implementation of corrective measures.

### 3. SOME RISKS OF DATA-DRIVEN INNOVATION IN ENGINEERING EDUCATION

The collection and processing of data for decision-making in engineering education is not without problems. In fact, there are important risks to worry about, which have been highlighted by numerous experts on learning analytics, educational data mining, and data protection policies, (Dringus, 2012) (Khalil, Taraghi, & Ebner, 2016), as well as certain ethical considerations (Slade, & Tait, 2019). Some of these risks are briefly described in this section, although each of them would deserve an entire paper for discussion.

A first problem refers to the secure storage of collected data. Many institutional education systems are not prepared to store large amounts of data in a secure way and are likely to have breaches that can compromise important data, including students' personal data. Relying on third-party services for data storage in the cloud can also lead to an inappropriate use of the data collected by these service providers.

A second problem refers to data privacy. It is very important to have an institutional strategy that clearly states what data needs to be collected from students and who has access to the data. Many educational institutions are not aware of all the data they collect (or can potentially collect); do not have mechanisms to control who has access to the data collected; and/or do not have an internal policy to facilitate that the right people can make proper informed decisions, being able to access the data they are entitled to.

A third problem refers to always getting explicit consent from the students for collecting data. It is very important that students (who actually own the data) know at all times what data are being collected from them and for what purpose the data are going to be used. In addition, there are international laws such as the GDPR (General Data Protection Regulation) in Europe that require, among other things, the removal of data collected upon request by its owner. Many educational institutions do not have data collection ethics committees and are not prepared to delete collected data upon request.

A fourth problem refers to transparency, or rather lack of transparency, of many algorithms and systems which are private and whose code is not open. The lack of transparency prevents algorithms and systems from being audited to better understand how they work.

Even if one relies on third-party algorithms and systems to make informed decisions, it is important to know how the results and visualizations they provide were obtained.

A fifth problem refers to bias. Many artificial intelligence systems use data collected in the past to make their calculations and predictions. However, data collected in the past may have important biases, such as a gender imbalance, which could lead to promoting more male students versus female students or vice versa. In general, it is important to take minorities into account when using data from the past as input so that these are not penalized.

Finally, it is important to bear in mind that humans may misinterpret the results and visualizations obtained from the data processed. Sometimes there are people who deliberately twist the data to fit a pre-designed theory, instead of discarding this theory if data advise to do so. Actually, from a very large dataset, and taking only a subset of it, erroneous and unreproducible conclusions can be very easily reached.

#### 4. CONCLUSIONS

Innovation in engineering education must be informed by data. Teachers must be aware of their students' performance (individually and at the class level) to support those who need more help as well as offer top quality educational resources progressively improved according to the data collected. Students should define a curriculum adapted to their particular needs and develop their SRL skills. Institutions must create specialized units for data collection and processing, and adequately train their staff (including teachers) for a data management culture. Nevertheless, there are numerous risks to be aware of, and a trade-off is needed so that these risks do not slow down innovation in educational institutions.

#### REFERENCES

- Alario-Hoyos, C., Pérez-Sanagustín, M., Delgado-Kloos, C., Parada G., H. A., & Muñoz-Organero, M. (2014). Delving into participants' profiles and use of social tools in MOOCs. *IEEE Transactions on Learning Technologies*, 7(3), 260-266.
- Alario-Hoyos, C., Muñoz-Merino, P. J., Pérez-Sanagustín, M., Delgado Kloos, C., & Parada G., H. A. (2016). Who are the top contributors in a MOOC? Relating participants' performance and contributions. *Journal of Computer Assisted Learning*, 32(3), 232-243.
- Alario-Hoyos, C., Estévez-Ayres, I., Pérez-Sanagustín, M., Delgado Kloos, C., & Fernández-Panadero, C. (2017). Understanding learners' motivation and learning *strategies in MOOCs. Intern. Review of Research in Open and Distributed Learning*, 18(3), 119-137.
- Alario-Hoyos, C., Estévez-Ayres, I., Delgado Kloos, C., Villena-Román, J., Muñoz-Merino, P. J., & Llorente-Pérez, E. (2019). Redesigning a freshman engineering course to promote

- active learning by flipping the classroom through the reuse of MOOCs. *International Journal of Engineering Education*, 35(1), 385-396.
- Alonso-Mencia, M. E., Alario-Hoyos, C., Maldonado-Mahauad, J., Estévez-Ayres, I., Pérez-Sanagustín, M., & Delgado Kloos, C. (2019). Self-regulated learning in MOOCs: lessons learned from a literature review. *Educational Review* (published online), 1-27.
- Baepler, P., Walker, J. D., & Driessen, M. (2014). It's not about seat time: Blending, flipping, and efficiency in active learning classrooms. *Computers & Education*, 78, 227-236.
- Delgado Kloos, C., Catalán-Aguirre, C., Muñoz-Merino, P. J., Alario-Hoyos, C. (2018). Design of a Conversational Agent as an Educational Tool, In *Learning with MOOCs 2018 (LWMOOCS V)* (pp. 27-30). IEEE.
- Dringus, L. P. (2012). Learning analytics considered harmful. *Journal of Asynchronous Learning Networks*, 16(3), 87-100.
- Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5/6), 304-317.
- Khalil, M., Taraghi, B., & Ebner, M. (2016). Engaging Learning Analytics in MOOCs: the good, the bad, and the ugly. In *Proceedings of the International Conference on Education and New Developments (END)* (pp. 3-7).
- Moreno-Marcos, P. M., Alario-Hoyos, C., Muñoz-Merino, P. J., & Delgado Kloos, C. (2018). Prediction in MOOCs: A review and future research directions. *IEEE Transactions on Learning Technologies* (published online), 1-19.
- Moreno-Marcos, P. M., Alario-Hoyos, C., Muñoz-Merino, P. J., Estévez-Ayres, I., & Delgado Kloos, C. (2018). Sentiment Analysis in MOOCs: A case study. In *2018 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1489-1496). IEEE.
- Pérez-Sanagustín, M., Hilliger, I., Alario-Hoyos, C., Delgado Kloos, C., & Rayyan, S. (2017). H-MOOC framework: reusing MOOCs for hybrid education. *Journal of Computing in Higher Education*, 29(1), 47-64.
- Ruipérez-Valiente, J. A., Muñoz-Merino, P. J., Leony, D., & Delgado Kloos, C. (2015). ALASKA: A learning analytics extension for better understanding the learning process in the Khan Academy platform. *Computers in Human Behavior*, 47, 139-148.
- Slade, S., & Tait, A. (2019). *Global guidelines: Ethics in learning analytics*. Oslo: International Council for Open and Distance Education (ICDE).
- Zimmerman, B. J. (2013). Theories of self-regulated learning and academic achievement: An overview and analysis. In *Self-regulated learning and academic achievement* (pp. 10-45). Routledge.

# Building Blocks for Powerful Ideas: Designing a Programming Language to Teach the Beauty and Joy of Computing

Jens Mönig  
jens.moenig@sap.com/ Research Expert SAP, Germany

Recepción: 9-8-2019 / Aceptación: 21-8-2019

**ABSTRACT.** Snap! is a cloud-native graphical programming environment and an online community. It is the programming language made for UC Berkeley’s popular introductory CS course named “The Beauty and Joy of Computing”. Snap! is taught in colleges and high schools across the U.S. from Palo Alto to Philadelphia. It has been translated to more than 40 languages and is used around the world—from Göttingen to Beijing—for teaching and research. Snap! has been designed for inclusion. Its low floor welcomes beginners and its multi-media capabilities invite creative thinkers of all ages. At the same time, Snap! offers sophisticated abstractions that make it suitable for an intellectually rigorous introduction to computer science.

**KEYWORDS:** Snap!, BJC, AP CSP, CS0.

## Construyendo bases sólidas para ideas poderosas: diseñando un lenguaje de programación para enseñar la belleza y alegría de la informática

**RESUMEN.** Snap! es un entorno de programación gráfica nativo de la nube y una comunidad en línea. Es el lenguaje de programación creado para el popular curso introductorio de CS de UC Berkeley llamado “La belleza y la alegría de la informática”. Snap! se imparte en colegios y escuelas secundarias de los EE. UU., desde Palo Alto hasta Filadelfia. Se ha traducido a más de 40 idiomas y se utiliza en todo el mundo, desde Gotinga hasta Beijing, para la enseñanza y la investigación. Snap! ha sido diseñado para su inclusión. El nivel bajo le da la bienvenida a principiantes y sus capacidades multimedia que invitan a pensadores creativos de todas las edades. Al mismo tiempo Snap! ofrece abstracciones sofisticadas que lo hacen adecuado para una introducción intelectualmente rigurosa a la informática.

**PALABRAS CLAVE:** Snap!, BJC, AP CSP, CS0.

## 1. INTRODUCTION

Recent years have seen a thunderous revival of programming education, sparked by a growing demand for computationally skilled workforce and spearheaded by MIT's visual Scratch language. In Scratch's wake, a new class of so-called "blocks-based" programming editors has appeared, and visual coding has since evolved into the de-facto standard for introductory CS activities. Along with Scratch's metaphor of stacking bricks, representing chunks of code into program-"towers" that are executed from top to bottom, a very traditional imperative style of programming has been established as quasi-best practice for introducing children and novices to CS.

At the same time, driven by the asynchronous nature of web programming and massive parallelization on the backend side to cope with "big data", many professional text-based programming languages have been revamped to support functional programming techniques such as proper tail-calls and even lambda, which before were considered too exotic to become mainstream. The gap between what beginners are exposed to in visual blocks-based languages and what is required to express themselves in a professional modern programming language today is more than just syntax. The gap is also conceptual and calls for proficiency in paradigms.

## 2. A CHALLENGE

Frequently discussed among educators is how to foster the transition from visual to textual programming. Sometimes the proposed solutions suggest bi-modal code editing, being able to switch back and forth between blocks and text. While this might address the lesser issue of coping with textual syntax, it does not help with introducing concepts and paradigms unsupported by any one side, and in the worst case even impoverishes the beginner's programming experience to the least common denominator of two programming languages.

On the other side of the spectrum, efforts are under way to broaden the scope and raise the ceiling of blocks-based programming. I will present one such project: Snap! Build Your Own Blocks. Snap! is a Scratch-like programming language that treats code blocks as first-class citizens instead of confining them to an editing modality. Embracing nested data structures and higher order functions, Snap! lets learners create arbitrary control structures and even custom programming languages with just blocks. Snap! has been developed for UC Berkeley's introductory computer science course named "The Beauty and Joy of Computing".

## 3. IN THIS TALK

I will share thoughts on the design of Snap! in a live-programmed excursion touching on a selection of powerful ideas from algorithms to artificial intelligence.

# PONENCIAS





# Análisis de sentimientos de noticias escritas usando un modelo basado en la red neuronal *long short-term memory* para determinar si las noticias positivas mejoran el estado de ánimo de las personas

Gustavo Adolfo Reyes-Paredes  
kivada@icloud.com / Universidad de Lima, Perú

Recepción: 17-6-2019 / Aceptación: 8-8-2019

RESUMEN. Es un hecho que el paradigma de distribuir noticias negativas a la población es el más aceptado mundialmente. Una gran cantidad de investigaciones se han enfocado en establecer los efectos de este paradigma en la población y, en todos los casos, se ha demostrado que es dañino para la salud y el comportamiento de las personas. Por ello, se ha decidido demostrar que el paradigma opuesto, la distribución de noticias positivas, genera una mejora en la salud, en el comportamiento y en el estado de ánimo de la población. Para lograr este propósito, se desarrolló un modelo basado en la red neuronal *long short-term memory* para realizar el análisis de los sentimientos relacionados con las noticias escritas en español. El experimento consistió en determinar el estado de ánimo de las personas luego de haber leído noticias positivas.

PALABRAS CLAVE: aprendizaje de máquinas, análisis de sentimientos, red neuronal recurrente, *long short-term memory*, bienestar psicológico y social

## Sentiment Analysis of Written News Using a Model Based on the Long Short-Term Memory Neural Network to Determine if Positive News Improve People's Mood

ABSTRACT. It is a fact that the paradigm of distributing negative news to the population is the most accepted worldwide. A large amount of research has been done to determine the effects of this paradigm on the population and, in all cases, it has been shown to be harmful to the health and behavior of people. Therefore, this paper aims to demonstrate that the opposite paradigm, the distribution of positive news, generates an improvement in the health, behavior and mood of the population. To achieve this, a model based on the long short-term memory neural network has been developed in order to analyze sentiments caused by news written in Spanish. Moreover, an experiment was conducted to determine people's mood after having read positive news.

KEYWORDS: machine learning, sentiment analysis, recurrent neural network, long short-term memory, psychological and social well-being

## 1. INTRODUCCIÓN

En la actualidad, el paradigma de difundir noticias negativas es el más aceptado, por ello la gran mayoría de los medios de comunicación las muestran a la población. Se ha demostrado que las personas se sienten atraídas hacia los eventos negativos en un nivel genético (Hatemi *et al.*, 2009). Por ello, no es sorprendente que la prensa mundial siga el paradigma de difundir noticias negativas, ya que esto genera un mayor ingreso monetario.

Por otro lado, es de suma importancia mantener un balance emocional positivo; de no lograrse, ocasiona una gran cantidad de problemas de salud afectando la vida diaria de las personas. La incapacidad de poder alcanzar un balance emocional es un problema actual y creciente que repercute en todo el mundo (Fredrickson, 2001); agravándose aún más por la proliferación de noticias negativas en los medios de comunicación, debido a que estas afectan la salud mental de la población si se exponen constantemente a ellas (Johnston y Davey, 1997).

Para reducir estos problemas existe otro paradigma basado en la exposición a la población de noticias positivas que nos indica que el estado de ánimo de las personas mejora luego de estar en contacto con noticias positivas. Por lo tanto, para corroborar estas investigaciones se propone realizar un modelo basado en la red neuronal recurrente *long short-term memory* (LSTM) que tiene como objetivo clasificar noticias escritas en positivas o negativas y luego realizar un experimento para determinar si el estado de ánimo de las personas realmente mejora.

Para ello se realizará una serie de tareas, las cuales resumidas son las siguientes:

- a) Recolectar una base de datos de noticias escritas en español.
- b) Clasificar estas noticias utilizando el servicio de Google AutoML, esto sirve para el entrenamiento.
- c) Transformar las noticias en vectores para que sirvan de *input* para el modelo.
- d) Formular e implementar el modelo de redes neuronales *long short-term memory*.
- e) Realizar la validación del modelo.
- f) Realizar el experimento para determinar el estado de ánimo de las personas luego de leer las noticias.

## 2. METODOLOGÍA

### 2.1 Recolección y preparación de noticias

Se utilizó una librería del lenguaje de programación Python llamada BeautifulSoup para poder realizar un web *scraping* de diferentes páginas web de noticias del Perú (*RPP, El Comercio, La República y Exitosa*)

Se escogió el título y el cuerpo de la noticia como *input* para el modelo, y se seleccionaron noticias entre el 12 de agosto del 2018 al 11 de septiembre del 2018 (Easton y McColl, 2007). Luego fueron almacenadas en un formato separado por comas (csv).

Para clasificar las noticias en negativas o positivas, primero fueron traducidas al inglés y luego se usó el servicio AutoML de Google para clasificarlas. Las noticias fueron traducidas al inglés debido a que AutoML de Google no puede determinar el sentimiento de textos en español.

En el siguiente paso, se empezó a realizar un proceso de muestreo aleatorio para balancear la proporción de noticias en 1:1, esto se realizó para que no exista un bias a la hora de entrenar el modelo. Al concluir este proceso se obtuvo una base de datos de noticias, de 20 000 noticias (10 000 noticias negativas y 10 000 noticias positivas) (Trochim, 2007).

Tabla 1  
*Distribución de las noticias*

Base de datos de noticias		
Fuente	Positivas	Negativas
<i>El Comercio</i>	2563	1709
<i>La República</i>	2896	2108
<i>RPP</i>	2563	3156
<i>Exitosa</i>	1978	3027
	10 000	10 000

Elaboración propia

Luego, se realizó un proceso de *encoding* y tokenización (Famili, Shen, Weber y Simoudis, 1997). De esta forma se transformarán las noticias en vectores de números. También se creó automáticamente un diccionario de palabras en el cual se identifican las palabras con un valor numérico, generado después del *encoding*.

En la tarea de tokenización se realizó la eliminación de *stop-words*, la eliminación de caracteres especiales y signos de puntuación (Klevecka y Lelis, 2008).

## 2.2 Desarrollo del modelo

Para determinar la polaridad de las noticias se decidió usar una granularidad a nivel del documento, esto se debe a que una noticia posee un texto extenso, entonces es necesario poder obtener la polaridad de este en forma conjunta, pues a lo largo del escrito puede haber diferentes puntos de vista, tomar la granularidad a nivel de oración podría haber sido perjudicial para la tarea de clasificación.

El modelo de aprendizaje de máquinas seleccionado fue de redes neuronales recurrentes, en específico el tipo LSTM (Hochreiter y Schmidhuber, 1997), este tipo se escogió debido a la capacidad de poder “recordar” los elementos de un texto, esto es fundamental en el análisis de textos largos porque de esta manera se puede capturar su contexto completo.

Para normalizar los datos de entrada (*inputs*) del modelo se usó la técnica de *mean pooling* por dos motivos: 1) para controlar la cantidad de *features* que van a ser recibidas por la capa de regresión logística; 2) para extraer la información promedio de cada texto, tomando en cuenta toda la información del texto, esto quiere decir que todos los valores son usados para realizar un mapeado de *features*.

Para determinar la clasificación de las noticias, en la capa final de la red neuronal, se usó una regresión logística, debido a que la salida de la red LSTM normalizada por el *average pooling* brinda valores entre 0 y 1, lo cual demuestra una probabilidad de ser positivos o negativos, pero para ser aún más exactos, se decidió usar la regresión logística ya que sólo se cuenta con dos posibles clasificaciones, positivas o negativas.

Se decidió recolectar un *dataset* propio de noticias de medios de comunicación locales, debido a que no existe uno. Igualmente, se recurrió al uso de técnicas de limpiado de datos como la eliminación de signos de puntuación y de *stopwords*, ya que ambos no agregan valor para la tarea a realizar. Luego se realizó un muestreo aleatorio en la base de datos para tenerla en una relación de 1:1 con respecto a noticias positivas y negativas, esto con el fin de que no exista un bias a la hora de entrenar el modelo.

El modelo propuesto es una variación de la red neuronal LSTM. En esta variación, la activación que ocurre en la puerta de salida de la capa LSTM no depende de su estado  $C_t$ , esto permite realizar parte de las operaciones necesarias en paralelo, volviendo el modelo mucho más eficiente computacionalmente, menos tiempo en entrenar el modelo sin algún impacto notable en la efectividad del modelo (Gers, Schmidhuber y Cummins, 2000).

Las ecuaciones siguientes describen como se actualiza la capa LSTM en cada unidad de tiempo  $t$ , según el modelo propuesto.

- 1) Es el *input* (datos de entrada) de la capa LSTM en un tiempo  $t$ .
- 2)  $W_f, W_c, W_o, U_i, U_f, U_c, U_o$  y  $V_o$  son matrices de *weights* (pesos del modelo).
- 3)  $b_i, b_f, b_c$  y  $b_o$  son vectores “bias”.

Primero, se calculan los valores para la puerta de entrada y el valor candidato para los estados de la capa LSTM en un tiempo  $t$ .

$$i_t = (W_i x_t + U_i h_{t-1} + b) \quad (1)$$

Donde  $i_t$  es el valor de la puerta de entrada en un tiempo  $t$  y  $\sigma$  es la función de activación.

$$\tilde{C}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \quad (2)$$

Donde  $\tilde{C}_t$  es el valor candidato para los estados de la red LSTM en un tiempo  $t$ .

Luego se calcula el valor de la función de activación  $f_t$  de las puertas de olvido en un tiempo  $t$ .

$$f_t = (W_f x_t + U_f h_{t-1} + b_f) \quad (3)$$

Luego de obtener el valor de la función de activación de la puerta de entrada, el valor de la función de activación de la puerta de salida y el valor candidato para los estados de la capa LSTM, se pasa a calcular el nuevo estado  $C_t$  de la capa LSTM en un tiempo  $t$ .

$$C_t = i_t * \tilde{C}_t + f_t * C_{t-1} \quad (4)$$

Donde  $\tilde{C}_t$  es el valor candidato para los estados de la red LSTM en un tiempo  $t$ .

Una vez calculado el nuevo estado  $C_t$  de la capa LSTM, se puede obtener el valor de sus puertas de salida y, como consecuencia, la salida final  $h_t$  de la capa LSTM.

$$o_t = (W_o x_t + U_o h_{t-1} + b_o) \quad (5)$$

Donde  $o_t$  es el valor de la puerta de salida de la red LSTM y  $\sigma$  es la función de activación.

$$h_t = o_t * \tanh(C_t) \quad (6)$$

La estructura final del modelo a usar consiste en una sola capa de redes neuronales LSTM, luego se implementará una capa de normalización (*mean pooling*), esto va a disminuir la variancia entre los valores, debido a que se va a tomar un promedio de la cantidad de datos para realizar un mapeo final de *features*. Después se implementará una capa de regresión logística para obtener una mayor eficacia a la hora de clasificar las noticias.

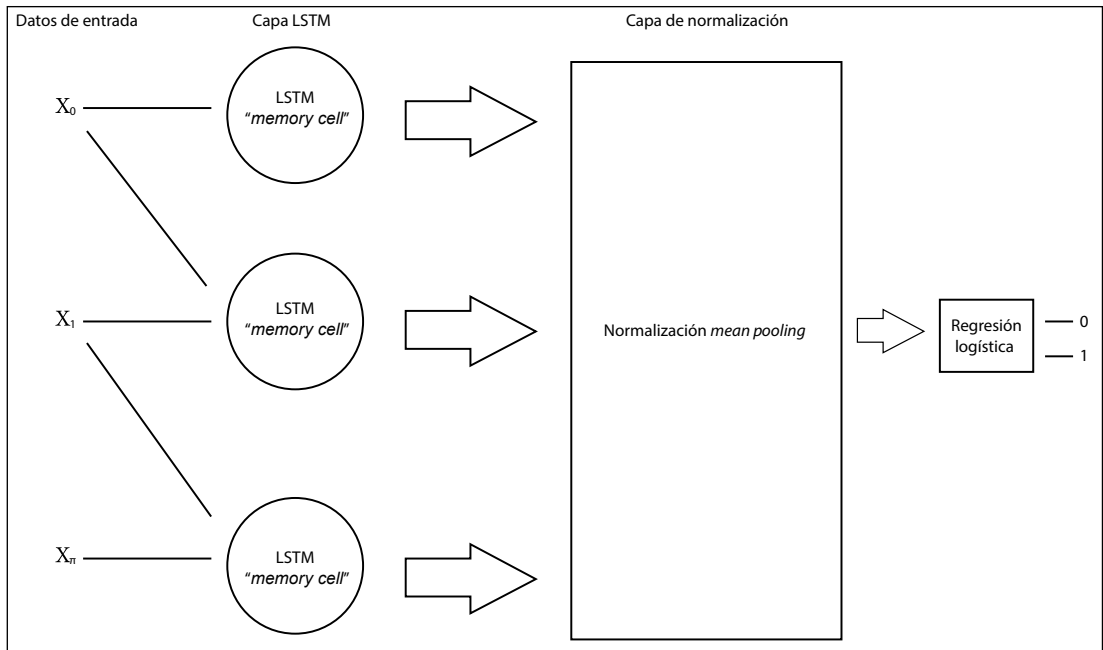


Figura 1. Estructura del modelo propuesto  
Elaboración propia

### 2.3 Algoritmo de optimización

Para incrementar la eficiencia de la red neuronal propuesta, se ha decidido implementar un algoritmo llamado Adam, el cual sirve para escoger el mejor *learning rate* para que el entrenamiento de la red neuronal sea el más rápido y efectivo posible (Kingma *et al.*, 2014).

Adam es un algoritmo de *learning rate* adaptativo, lo que significa que calcula *learning rates* individuales para diferentes parámetros. Su nombre se deriva de la estimación del momento adaptativo, y la razón por la que se llama así es porque Adam usa estimaciones del primer y segundo momento del gradiente para adaptar la velocidad de aprendizaje para cada *weight* de la red neuronal.

Adam utiliza los gradientes elevados al cuadrado para escalar la velocidad de aprendizaje como el algoritmo RMSprop (Tieleman y Hinton, 2012) y aprovecha el impulso al usar la media móvil de los gradientes en lugar de solo la gradiente como el algoritmo *stochastic gradient descent*.

Para estimar los momentos, Adam utiliza promedios móviles exponenciales, calculados en el gradiente evaluado en un minilote actual:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (7)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (8)$$

Donde  $m_t$  y  $v_t$  son los promedios móviles,  $g$  es la gradiente del minilote actual y los  $\beta$  son los hiperparámetros del algoritmo. Ambos tienen un valor *default* de 0,9 y 0,999, respectivamente

La fórmula para obtener el promedio móvil  $m_t$  también puede ser representada por la siguiente ecuación:

$$m_t = (1 - \beta_1) \sum_{i=0}^t \beta_1^{t-i} g_i \quad (9)$$

Después de esto se necesita realizar un paso llamado “*bias correction*”, esto quiere decir que se necesita corregir el valor inicial de  $m_t$  y  $v_t$ , para ello se realizan las siguientes transformaciones:

$$\hat{m}_t = 1^m - \beta_1^t \quad (10)$$

$$\hat{v}_t = 1 - v_t \beta_2 \quad (11)$$

Para finalizar, se utilizan estos promedios móviles  $\hat{m}_t$  y  $\hat{v}_t$  para escalar la *learning rate* individualmente para cada parámetro. La forma en que se hace en Adam es simple, para realizar una actualización de *weights* hacemos lo siguiente:

$$W_t = W_{t-1} - \eta \left( \sqrt{\hat{v}_t} + \varepsilon \right) \quad (12)$$

Donde  $W_t$  es una matriz *weights* y la  $\eta$  es el tamaño del *step* (Zhang, Ma, Li y Wu, 2017).

## 2.4 Impacto de las noticias positivas

Para poder validar la hipótesis se decidió usar el test PANAS (Positive and Negative Affect Schedule).

Fue desarrollado en 1988 por Watson, Clark y Tellegen con el objetivo de medir de una forma más pura el efecto negativo o positivo de algún evento en específico (Watson, Clark y Tellegen, 1988). En un inicio el test PANAS contaba con 60 ítems (sentimientos) para determinar el sentimiento negativo o positivo, pero luego se realizó una reducción de la lista, basada en el coeficiente de importancia, y al final se obtuvieron 20 ítems, 10 para el sentimiento negativo y 10 para el sentimiento positivo (Watson, Clark y Tellegen, 1988).

Los resultados de este cuestionario fueron alentadores, porque fueron acertados, presentando los siguientes porcentajes: para la escala de sentimiento positivo, el coeficiente alfa de Cronbach (Cronbach, 1951) fue de 0,86 a 0,90; para la escala de sentimiento negativo, 0,84 a 0,87 (Watson, Clark y Tellegen, 1988).

### 3. RESULTADOS

#### 3.1 Validación del modelo LSTM

Luego de haber implementado el modelo, se probó su eficacia utilizando el método de validación cruzada *k-folds* con un  $k=10$  (Kohavi, 1995). Los resultados fueron los siguientes:

Tabla 2  
*Resultados de la validación cruzada k-folds*

Validación cruzada <i>k-folds</i> (k=10)	
Folds (k)	Accuracy
1	86,56 %
2	85,36 %
3	86,85 %
4	89,37 %
5	88,29 %
6	85,67 %
7	89,95 %
8	91,51 %
9	88,30 %
10	87,92 %
Promedio: 87,98 %	

Elaboración propia

Para comparar la efectividad del modelo se implementaron otros dos métodos que también son usados en la actualidad:

- 1) Naive Bayes
- 2) Red neuronal recurrente (RNN)

El método usado para la validación, por consistencia, fue el de validación cruzada *k-folds* ( $k=10$ ).



1) *Naive Bayes*

Tabla 3  
*Resultados de la validación cruzada k-folds*

Validación cruzada <i>k-folds</i> (k=10)	
Folds (k)	Accuracy
1	79,89 %
2	83,25 %
3	79,28 %
4	81,59 %
5	83,17 %
6	79,03 %
7	79,98 %
8	80,55 %
9	81,71 %
10	82,99 %
Promedio: 81,14 %	

Elaboración propia

2) *Red neuronal recurrente RNN*

Tabla 4  
*Resultados de la validación cruzada k-folds*

Validación cruzada <i>k-folds</i> (k=10)	
Folds (k)	Accuracy
1	82,79 %
2	81,90 %
3	83,58 %
4	81,27 %
5	82,90 %
6	83,51 %
7	84,95 %
8	80,11 %
9	82,76 %
10	80,56 %
Promedio: 82,43 %	

Elaboración propia

Luego de ver estos resultados, se puede afirmar que el modelo propuesto, una variación de la red neuronal recurrente LSTM tiene un mejor *accuracy* que el método de Naive Bayes y la red neuronal recurrente.

### 3.2 Procedimiento experimental

#### *Test PANAS*

Para validar la hipótesis se realizó un experimento basado en encuestas, con la participación de 520 personas de edades diferentes (entre 20-45 años) y de diferentes géneros (320 mujeres, 200 hombres). La encuesta se realizó antes y después de haber leído las noticias. La muestra se dividió de la siguiente forma:

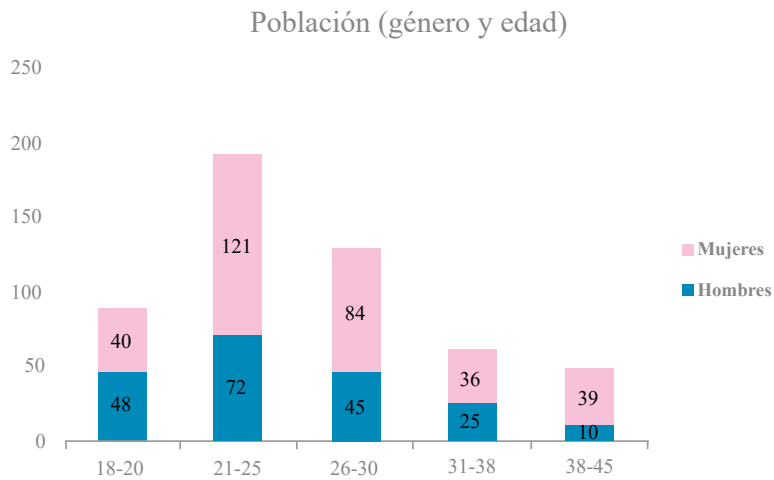


Figura 2. Perfil demográfico (género y edad)  
Elaboración propia

Después de haber realizado las encuestas, los resultados fueron los siguientes:

Tabla 5  
Resultados del test PANAS, antes y después de haber leído las noticias positivas

		Resultados Test PANAS			
		Positivo		Negativo	
		Antes	Después	Antes	Después
Hombres	18-20	25	39	30	16
	21-25	19	40	38	18
	26-30	15	38	42	20
	31-38	12	38	40	16
	38-45	14	40	41	17
Mujeres	18-20	27	42	25	12
	21-25	20	45	31	14
	26-30	12	37	38	17
	31-38	15	38	40	10
	38-45	17	42	38	19

Elaboración propia

#### 4. CONCLUSIONES

Luego de haber realizado este reporte, implementando tanto el modelo de LSTM, para poder clasificar las noticias escritas en español en positivas o negativas, como el test de PANAS, con una población variada para poder determinar el efecto de las noticias positivas en ellas, se pueden concluir dos grandes puntos:

- a) El modelo propuesto basado en las redes neuronales LSTM cuenta con un *accuracy* del 87,98 %, esto es bueno comparado con el método de Naive Bayes y la red neuronal recurrente.
- b) Los resultados del test de PANAS nos demuestran que todas las personas que han leído las noticias positivas han mejorado su estado de ánimo.

#### 5. TRABAJOS FUTUROS

Con respecto a la primera parte de este artículo, existen otras técnicas que pueden ser utilizadas a cambio de las propuestas en este artículo, como *max-pooling* o una versión del algoritmo Adam (NDAAdam) que resuelve alguno de sus errores (Zhang *et al.*, 2017).

También existen otros métodos, en especial el enfoque basado en la granularidad de aspecto, en contraste con la granularidad de texto, usada en este artículo, tiene un gran campo de investigación abierto, ya que usando este enfoque se pueden reconocer diferentes puntos de vista que existen en un texto, así como también el sarcasmo.

Por otro lado, las técnicas basadas en aprendizaje de máquinas no supervisadas podrían representar el futuro en este campo de investigación, esto se debe a la inmensa cantidad de datos que se generan diariamente y que en su mayoría no se encuentran clasificados.

## REFERENCIAS

- Bastien, F., Lamblin, P., Pascanu, R., Bergstra, J., Goodfellow, I., Bergeron, A., Bouchard, N., Warde-Farley, D., y Bengio, Y. (2012). Theano: new features and speed improvements. *Deep Learning Workshop, NIPS 2012*.
- Bergstra, J., Breuleux, O., Bastien, F., Lamblin, P., Pascanu, R., Desjardins, G., Turian, J. P., Warde-Farley, D., y Bengio, Y. (2010). Theano: a CPU and GPU math expression compiler in Python. *Proceedings of the Python for Scientific Computing Conference (SciPy 2010)*.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334. doi:10.1007/BF02310555.
- Easton, V. J., y McColl, J. H. (2007). Statistics glossary: Sampling. Recuperado de <http://www.stats.gla.ac.uk/steps/glossary/sampling.html>
- Famili, A., Shen, W.-M., Weber, R., y Simoudis, E. (1997). Data preprocessing and intelligent data analysis. *Intelligent Data Analysis*, 1(1), 3-23.
- Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist*, 56(3), 218-226.
- Gers, F., Schraudolph, N. N., y Schmidhuber, J. (2002). Learning precise timing with LSTM recurrent networks. *Journal of Machine Learning Research*, 3(1), 115-143.
- Gers, F. A., Schmidhuber, J., y Cummins, F. A. (2000). Learning to forget: Continual prediction with LSTM. *Neural Computation*, 12(10), 2451-2471.
- Hatemi, P. K., Funk, C. L., Medland, S. E., Maes, H. M., Silberg, J. L., Martin, N. G., y Eaves, L. J. (2009). Genetic and Environmental Transmission of Political Attitudes Over a Life Time. *Journal of Politics*, 71(3), 1141-1156.
- Hochreiter, S., y Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735-1780.

- Kingma, D. P., y Ba, J. (2015). Adam: A method for stochastic optimization. CoRR, abs/1412.6980.
- Klevecka, I., y Lelis, J. (2008). Pre-processing of input data of neural networks: The case of forecasting telecommunication network traffic. *Teletronikk: Telecommunications Forecasting (Special issue in co-operation with International Institute of Forecasters)*, 104(3/4): 168-178.
- Johnston, W. M., y Davey, G. C. (1997). The psychological impact of negative TV news bulletins: The catastrophizing of personal worries. *British Journal of Psychology*, 88, 85-91. doi:10.1111/j.20448295.1997.tb02622.x
- Tieleman, T., y Hinton, G. (2012). Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude. COURSERA: *Neural Networks for Machine Learning*, 4(2), 26-31,
- Trochim, W. K. (2007). Sampling. Research methods knowledge base. Recuperado de <http://www.socialresearchmethods.net/kb/sampling.php>
- Watson, D., Clark, L. A., y Tellegen, A. (1988). Development and validation of brief measures of positive and neative affect: The PANAS Scales. *Journal of Personality and Social Psychology*, 54(6), 1063-1070. doi:10.1037/0022-3514.54.6.1063
- Zhang, Z., Ma, L., Li, Z., y Wu, C. (2018). Normalized direction-preserving Adam. arXiv:1709.04546v2



# Modelo de selección de plataforma educativa virtual con mapas cognitivos difusos (FCM)

Manuela Linares-Barbero  
malinare@ulima.edu.pe/ Universidad de Lima, Perú

Recepción: 30-5-2019 / Aceptación: 9-7-2019

**RESUMEN.** La tendencia mundial contemporánea en educación se inclina hacia la formación virtual, la cual ofrece mayor versatilidad en lo referente a la educación a distancia, pues cuenta con numerosas herramientas al alcance de los estudiantes. Los cursos en línea, en auge, especialmente los de modalidad abierta en línea y masivos (*massive online open course* - MOOC), han logrado difundir el conocimiento a la mayor cantidad de alumnos en poco tiempo, fenómeno que antes no era posible. Al iniciarse la transformación hacia la educación virtual, es indispensable el uso de una plataforma virtual. La decisión sobre esta determina las herramientas que se tengan disponibles, así como las posibilidades de ofrecer cursos semipresenciales, cursos 100 % en línea, cursos presenciales con talleres virtuales o cursos abiertos masivos en línea (MOOC). Por tal razón, hemos descubierto la importancia del desarrollo de un modelo original para la selección de una plataforma virtual educativa. Para el propósito de esta investigación elaboraremos el mapa cognitivo difuso con dos rondas de entrevistas a expertos en el tema.

**PALABRAS CLAVE:** mapas cognitivos difusos, FCM, plataforma virtual, toma de decisiones, educación superior, *e-learning*, educación a distancia, cursos abiertos masivos en línea, MOOC, sistemas de información

## Virtual Educational Platform Selection Model Using Fuzzy Cognitive Maps (FCM)

**ABSTRACT.** Contemporary global trends in education favor virtual training, which offers greater versatility in terms of distance education, as it has numerous tools available for students. Online courses, currently on the rise, especially massive open online courses (MOOCs), have managed to spread knowledge to the largest number of students in a short time: a phenomenon that was not possible before. When transformation towards virtual education begins, using a virtual platform is essential. Decisions on this determine the tools that will be available, as well as the possibilities of offering semi-face-to-face courses, 100 % online courses, face-to-face courses with virtual workshops or MOOCs. For this reason, we have discovered the importance of developing an original model for selecting a virtual educational platform. For the purpose of our research, we will prepare a fuzzy cognitive map with two rounds of interviews with experts in the field.

**KEYWORDS:** fuzzy cognitive maps, FCM, virtual platform, decision-making, higher education, e-learning, distance learning, massive open online courses, MOOC, information systems

## 1. INTRODUCCIÓN

El proceso de enseñanza evoluciona conforme a los avances tecnológicos, superando las barreras geográficas y ofreciendo a la sociedad abundantes conocimientos a través de diferentes tipos de contenidos. Sin embargo, la propuesta de educación a distancia ni es nueva ni exclusiva de las plataformas virtuales, por ejemplo, el caso de Open University, fundada a fines de la década de 1960 con el propósito de promover aprendizajes a través de una universidad de educación a distancia (Issroff y Eisenstadt, 1997). Desde entonces, muchas modalidades de estudios a distancia o por correspondencia existen y se presentan como una alternativa para quienes no pueden asistir presencialmente a un curso. Tal es el caso de la escuela de verano virtual de la Open University, que en el verano de 1994 impartió un curso piloto con doce alumnos, quienes contaron con una Apple Macintosh, un módem y un celular<sup>1</sup> (Issroff y Eisenstadt, 1997). De momento existe una variedad de páginas web que ofrecen cursos en línea masivos y abiertos, (MOOC), como Udacity, EdX, MiríadaX, Coursera (Ramírez-Fernández y Salmerón Silvera, 2015), entre otros; y en universidades como Stanford y MIT (Ho, Ke y Liu, 2015).

Actualmente la mayoría de las universidades a nivel mundial cuenta con algún tipo de plataforma de soporte a la enseñanza, así sea utilizada solo como repositorio de los materiales del curso. Igualmente, advertimos grandes cambios que se proyectan respecto a la influencia que tiene la tecnología incrementando así la enseñanza virtual y dejando de lado las aulas presenciales. Uno de los motivos por el cual esta forma de aprendizaje es tan atractiva, como lo manifiestan las instituciones que la emplean, especialmente, las que ofrecen cursos MOOC, es la reducción de costos (Ho *et al.*, 2015; Tarhini, Hone y Liu, 2014), y el incremento de los ingresos (Tarhini *et al.*, 2014). De esta forma, descubrimos como los centros educativos alcanzan una serie de beneficios económicos a través del *e-learning* (Ho *et al.*, 2015).

Justamente este crecimiento exponencial de cursos virtuales nos lleva a detectar estándares de gestión de la calidad de formación virtual, como UNE 66181 de AENOR (Asociación Española de Normalización y Certificados) (Ramírez-Fernández y Salmerón Silvera, 2015). Una de las ventajas que se logra es el incremento del conocimiento (Ho *et al.*, 2015), que apoyado con las tecnologías de la comunicación, puede llegar a mayor cantidad de personas dada la flexibilidad geográfica y de tiempo para estudiar (Tarhini *et al.*, 2014). Aun así, podemos ver que la ratio de alumnos que se retira del curso es de seis a siete veces mayor en los cursos en línea en comparación con los presenciales (Tarhini *et al.*, 2014).

La decisión sobre la selección de una nueva plataforma normalmente recae en la gerencia o en los instructores que ofrecen el servicio. No siempre están involucrados los estudiantes y/o profesores (Ho *et al.*, 2015). Por este motivo es de vital importancia que se modele la selección de una plataforma virtual educativa que cuente con una herramienta útil que pueda ser utilizada por todos para una mejor calidad en el proceso de toma de decisiones.

---

1 La razón radica en que el módem ocuparía la línea de teléfono fija de la casa del alumno.



Presentamos en este artículo un modelo de mapa cognitivo difuso (*fuzzy cognitive map* - FCM) para la selección de plataformas educativas virtuales en universidades. En la segunda sección detallaremos la metodología elegida y como fue empleada. En la tercera parte mostraremos los resultados de la investigación y en la cuarta explicaremos las conclusiones basadas en los resultados obtenidos.

## 2. METODOLOGÍA

Los modelos de mapas cognitivos tienen sus orígenes con Axelrod (1976), quien, en busca de mejorar la calidad de vida de las personas, planteó el enfoque de mapas cognitivos para la toma de decisiones. Posteriormente, Kosko (1986) aportó a los mapas cognitivos el uso de lógica difusa y dio paso a los mapas cognitivos difusos. Este método se basa en la diagramación del proceso mental de toma de decisiones de las personas y facilita, como diagrama difuso direccional, la retroalimentación. Igualmente permite modificar los valores iniciales y lo hace dinámico (Kosko, 1988). El estado de las variables evoluciona en base a los valores difusos asignados en la relación (Salmerón y López, 2012). El mapa cognitivo contiene nodos que significan conceptos o variables. El mapa cognitivo difuso está representado por una matriz de  $n \times n$ , siendo  $n$  el número de nodos (Mirghafoori, Sharifabadi y Takalo, 2018). Si elaboramos un modelo del proceso de decisión de los expertos, podremos poner en manos del gran público la posibilidad de tomar la mejor decisión sobre un tema determinado, tal como lo hubiera hecho un especialista.

Los mapas cognitivos difusos consisten en sistemas neuronales difusos que son capaces de modelar sistemas complejos (Kosko, 1986). Se cuenta con tres elementos: los conceptos diagramados con nodos, los arcos, que representan el sentido de la relación de causa-efecto, y la intensidad que varía de  $-1$  a  $1$  (Tornese, Verriello, Gnoni, Mossa, y Mummolo, 2017; Pandari y Azar, 2017; Salmerón y López, 2012). Los mapas cognitivos difusos son usados con mucho éxito en la toma de decisiones en ambientes complejos, de gran incertidumbre y con poco tiempo para actuar (Salmerón y López, 2012).

Desde los descubrimientos de Axelrod, el método de aplicación de los mapas cognitivos para la toma de decisiones evolucionó, adicionando un elemento de la lógica difusa (*fuzzy logic*) y dando paso a los mapas cognitivos difusos (FCM). En estos últimos se aprecian varias modificaciones como los mapas cognitivos difusos basados en reglas o *rule-based fuzzy cognitive maps* (RBFCM), mapas cognitivos difusos grises, mapas cognitivos difusos intuicionistas, redes cognitivas dinámicas, mapas cognitivos difusos dinámicos aleatorios, redes cognitivas difusas, mapas cognitivos difusos evolutivos, mapas cognitivos difusos de tiempo, reglas difusas incorporadas a los mapas cognitivos difusos, mapas cognitivos difusos con grado de creencias distribuidos y mapas cognitivos en bruto o *rough cognitive maps* (Papageorgiou y Salmerón, 2013). Los mapas cognitivos difusos (FCM), en su forma original, toman en cuenta tres elementos: la dirección de la relación de causa-efecto entre los nodos, representada por una flecha; la intensidad de dicha

relación, representada por un valor numérico entre 0 y 1; y si la relación es directa o inversa, representada por un signo positivo si es directa y uno negativo si es inversa (Bueno y Salmerón, 2008). En este artículo utilizamos el mapa cognitivo difuso (FCM) en su forma original.

### 3. RESULTADOS

Se ha utilizado el método Delphi para la validación de los factores críticos en la selección de la plataforma virtual educativa, así como para determinar la relación de causa-efecto entre los factores y su intensidad, a través de dos rondas de entrevistas a expertos en educación superior. Estos se desempeñan, o lo han sido en el pasado, en cargos donde fueron partícipes de decisiones en selección de plataformas educativas o similares. Los mencionados especialistas ostentaron los siguientes puestos de trabajo: un vicerrector, tres directores, un subdirector, cuatro coordinadores y dos jefes de proyectos; en seis universidades, Universidad de Lima (ULima), Pontificia Universidad Católica del Perú (PUCP), Universidad Peruana de Ciencias Aplicadas (UPC), Universidad San Ignacio de Loyola (USIL), Universidad Tecnológica del Perú (UTP) y Universidad de Ciencias y Artes de América Latina (UCAL), y un instituto, Instituto Superior Tecnológico Privado de Comercio Exterior (ADEX). En total, fueron once expertos con un mínimo de seis años de experiencia profesional en el rubro de la educación superior.

Para la definición del número de entrevistados, dentro del método Delphi se suelen utilizar entre diez a quince conocedores, sin embargo, mientras mayor es la heterogeneidad del grupo, menor deben ser los expertos; entonces, es recomendable recurrir a un grupo entre cinco y diez sujetos (Clayton, 1997). Al tener grupos heterogéneos es importante contar con al menos un especialista de cada tipo. También es cierto que, mientras sea mayor la cantidad de personas en todas las rondas, se logrará mayor estabilidad en los resultados (Bueno y Salmerón, 2008). En esta investigación se reunieron un total de once expertos en la primera ronda y de diez (un vicerrector, dos directores, un subdirector, cuatro coordinadores y dos jefes de proyectos) en la segunda.

Luego de una extensa revisión de la literatura se identificaron los factores críticos en la selección de una plataforma virtual, divididos en tres dimensiones, la primera relacionada exclusivamente a la educación donde se identifican once factores críticos, la segunda relacionada al sistema donde se descubren once factores críticos y la tercera relacionada a la organización donde se detectan ocho factores críticos (véase la tabla 1).

En ambas rondas con los expertos se utilizó una escala Likert de 1 a 7, siendo 1 nada importante y 7 muy importante, para que evaluaran la importancia de los factores críticos mencionados. Además, se les preguntó si encontraban alguno adicional. Un grupo de expertos mencionó uno o dos adicionales, sin embargo, estos se eliminaron en la segunda ronda por tener baja importancia. En la tabla 1 se puede ver el resultado de la importancia de los factores mostrando la media aritmética de cada uno.

Tabla 1  
Factores críticos y su valoración de importancia

n.º	Factor	Media
Relacionados a la educación		
1	Organización y planificación del curso	6,55
2	Claridad en la comunicación, habilidades de comunicación	6,55
3	Interacción entre el profesor y el estudiante	6,36
4	Dificultad del curso, cantidad de trabajos	4,91
5	Evaluación y exámenes	5,55
6	Percepción de aprendizaje del estudiante	6,18
7	Diseño didáctico-instruccional	6,18
8	Recursos formativos y actividades de aprendizaje	6,36
9	Accesibilidad <i>hardware</i>	5,36
10	Accesibilidad <i>software</i>	5,91
11	Accesibilidad web	6,45
Relacionados al sistema		
12	Capacidad de integrar la plataforma virtual con el actual IS/IT	5,64
13	Modularidad	5,64
14	Confianza en la plataforma virtual	6,36
15	Adaptabilidad de la plataforma virtual a las necesidades actuales	6,00
16	Capacidad de la plataforma virtual de entregar información a tiempo	6,27
17	Costo del <i>software</i>	5,00
18	Costo de mantenimiento	5,00
19	Costo de la consultoría o implementación	4,91
20	Requerimientos de un equipo especializado	5,27
21	Servicios del <i>software</i>	5,45
22	Prestigio del <i>software</i>	5,27
Relacionados a la organización		
23	El personal realiza educación continua frecuentemente	5,27
24	Promedio de edad del personal	3,82
25	El grupo de decisión realiza educación continua frecuentemente	5,45
26	Cultura organizacional tradicional	4,27
27	Número de trabajadores / tamaño de la empresa	4,27
28	Estrategia organizacional tradicional	4,27
29	Complejidad de la organización	5,00
30	Capacidad de flexibilidad de la organización	5,64

Fuente: Elaboración propia con los factores de Ramírez-Fernández y Salmerón Silvera (2015), Rovai, Ponton, Derrick y Davis (2006) y Bueno y Salmerón (2008)

Se aprecia que los factores 24, 26, 27 y 28 muestran una media aritmética menor a 4,5 de los 7 puntos de la escala de Likert, esto indica que dichos factores no son los más importantes en la selección de una plataforma educativa virtual, razón por la cual se eliminaron del modelo. También se interrogaron a los expertos sobre la relación de causa-efecto entre los factores por dimensión, para identificar el sentido de esta relación, si es directa o inversa, así como la intensidad de la misma. En la tabla 2 se observa el sentido de la relación causa-efecto en la columna “Relación”, siendo el primer número el factor causante y el segundo número el factor afectado. Así mismo, se ven la media aritmética, la mediana, la moda, la desviación estándar, el valor mínimo y el valor máximo de la intensidad que genera esta relación de causa-efecto de acuerdo con los expertos.

Tabla 2  
Resultados de la intensidad de las relaciones

Relación	Media	Mediana	Moda	Desviación estándar	Valor mín.	Valor máx.
Relacionados a la educación						
1-2	0,88	1,00	1	0,20	0,50	1,00
1-3	0,75	0,90	1	0,39	-0,20	1,00
1-4	0,50	0,70	0,7	0,57	-0,80	1,00
1-5	0,68	0,80	1	0,53	-0,80	1,00
1-6	0,75	0,90	1	0,38	0,00	1,00
1-7	0,91	0,90	1	0,12	0,60	1,00
1-8	0,89	1,00	1	0,14	0,60	1,00
1-9	0,36	0,10	0,1	0,42	0,00	1,00
1-10	0,44	0,20	0	0,44	0,00	1,00
1-11	0,68	0,80	1	0,39	0,00	1,00
2-3	0,91	1,00	1	0,16	0,50	1,00
4-2	0,17	0,50	0,5	0,67	-1,00	1,00
2-5	0,56	0,80	0,8	0,58	-1,00	1,00
2-6	0,86	0,90	1	0,23	0,20	1,00
2-7	0,57	0,90	1	0,58	-0,80	1,00
2-8	0,74	0,90	1	0,34	0,10	1,00
2-9	0,28	0,10	0	0,37	0,00	1,00
2-10	0,44	0,20	0	0,44	0,00	1,00
2-11	0,61	0,80	1	0,44	0,00	1,00

(continúa)

(continuación)

4-3	0,28	0,50	1	0,76	-1,00	1,00
3-5	0,45	0,60	1	0,59	-1,00	1,00
3-6	0,82	0,90	1	0,26	0,20	1,00
7-3	0,63	0,70	1	0,40	0,00	1,00
8-3	0,64	0,70	1	0,39	0,00	1,00
3-9	0,30	0,20	0	0,49	-0,60	1,00
3-10	0,45	0,60	1	0,54	-0,60	1,00
3-11	0,55	0,80	1	0,54	-0,60	1,00
4-5	0,15	0,50	0,5	0,73	-1,00	1,00
4-6	0,51	0,60	1	0,61	-0,80	1,00
4-7	0,75	0,90	1	0,38	-0,10	1,00
4-8	0,61	0,90	1	0,59	-0,70	1,00
4-9	0,30	0,20	0	0,35	0,00	1,00
4-10	0,60	0,70	1	0,46	-0,30	1,00
4-11	0,65	0,70	1	0,39	0,00	1,00
5-6	0,68	0,90	1	0,39	0,00	1,00
5-7	0,66	0,60	1	0,36	0,00	1,00
5-8	0,70	0,80	1	0,34	0,10	1,00
5-9	0,39	0,20	0	0,40	0,00	1,00
5-10	0,67	0,80	1	0,37	0,00	1,00
5-11	0,66	0,80	1	0,38	0,00	1,00
7-6	0,85	0,90	1	0,17	0,50	1,00
6-8	0,85	0,90	1	0,19	0,50	1,00
6-9	0,55	0,70	1	0,41	0,00	1,00
6-10	0,60	0,70	0,7	0,40	0,00	1,00
6-11	0,73	0,90	1	0,37	0,00	1,00
7-8	0,90	1,00	1	0,16	0,60	1,00
7-9	0,51	0,50	1	0,39	0,00	1,00
7-10	0,66	0,70	1	0,36	0,00	1,00
7-11	0,68	0,70	1	0,36	0,00	1,00
8-9	0,62	0,70	1	0,40	0,00	1,00
8-10	0,77	0,90	1	0,32	0,00	1,00
8-11	0,77	0,90	1	0,32	0,00	1,00
9-10	0,79	1,00	1	0,30	0,10	1,00
9-11	0,79	1,00	1	0,30	0,10	1,00
10-11	0,83	1,00	1	0,23	0,50	1,00
11-31	0,76	0,70	0,7	0,11	0,70	1,00

(continúa)

(continuación)

Relacionados al sistema						
13-12	0,71	0,90	1	0,37	0,00	1,00
12-14	0,64	0,80	1	0,39	0,00	1,00
15-12	0,91	0,90	1	0,10	0,70	1,00
16-12	0,78	0,90	1	0,28	0,20	1,00
12-17	0,26	0,10	0,1	0,49	-0,60	1,00
12-18	0,25	0,10	0,1	0,51	-0,60	1,00
12-19	0,33	0,50	0,5	0,56	-0,60	1,00
12-20	0,35	0,50	1	0,65	-0,60	1,00
12-21	0,65	0,90	1	0,51	-0,60	1,00
12-22	0,65	0,50	1	0,38	0,00	1,00
13-14	0,48	0,50	1	0,43	0,00	1,00
13-15	0,69	0,90	1	0,37	0,00	1,00
13-16	0,60	0,80	1	0,41	0,00	1,00
13-17	0,44	0,30	0	0,43	0,00	1,00
13-18	0,39	0,10	0	0,45	0,00	1,00
13-19	0,43	0,50	0	0,42	0,00	1,00
13-20	0,30	0,30	1	0,59	-0,70	1,00
13-21	0,58	0,50	1	0,41	0,00	1,00
13-22	0,36	0,20	0	0,42	0,00	1,00
15-14	0,78	0,90	1	0,31	0,00	1,00
16-14	0,85	1,00	1	0,21	0,40	1,00
14-17	0,50	0,50	1	0,38	0,00	1,00
14-18	0,39	0,20	0,2	0,38	0,00	1,00
14-19	0,48	0,40	1	0,40	0,00	1,00
14-20	0,37	0,20	1	0,57	-0,60	1,00
14-21	0,45	0,20	1	0,60	-0,80	1,00
14-22	0,48	0,50	1	0,60	-0,80	1,00
15-16	0,89	1,00	1	0,19	0,40	1,00
15-17	0,45	0,50	0,5	0,37	0,00	1,00
15-18	0,43	0,20	0,2	0,38	0,00	1,00
15-19	0,56	0,60	1	0,37	0,10	1,00
15-20	0,38	0,20	1	0,47	-0,30	1,00
15-21	0,54	0,60	1	0,48	-0,20	1,00
15-22	0,59	0,60	1	0,41	0,00	1,00
16-17	0,51	0,50	0,5	0,36	0,00	1,00

(continúa)

(continuación)

16-18	0,38	0,20	0,1	0,40	0,00	1,00
16-19	0,50	0,40	0,1	0,38	0,10	1,00
16-20	0,39	0,40	1	0,51	-0,70	1,00
16-21	0,47	0,50	1	0,53	-0,70	1,00
16-22	0,49	0,40	1	0,38	0,00	1,00
17-18	0,49	0,50	1	0,47	-0,40	1,00
17-19	0,55	0,50	1	0,45	-0,40	1,00
17-20	0,55	0,50	1	0,40	0,10	1,00
17-21	0,68	0,70	1	0,33	0,10	1,00
17-22	0,59	0,60	1	0,37	0,00	1,00
18-19	0,42	0,40	0,2	0,36	0,00	1,00
18-20	0,50	0,30	0,2	0,39	0,10	1,00
18-21	0,60	0,50	1	0,35	0,20	1,00
18-22	0,32	0,20	0,2	0,43	-0,50	1,00
19-20	0,53	0,50	1	0,44	-0,30	1,00
19-21	0,71	0,70	1	0,30	0,20	1,00
19-22	0,49	0,50	1	0,43	-0,30	1,00
20-21	0,65	0,80	1	0,36	0,00	1,00
20-22	0,59	0,80	1	0,45	-0,30	1,00
21-22	0,68	1,00	1	0,42	0,00	1,00
22-31	0,56	0,50	0,5	0,12	0,50	0,85
Relacionados a la organización						
23-25	0,63	0,70	1	0,40	0,00	1,00
23-29	0,26	0,30	0,5	0,51	-0,60	1,00
23-30	0,74	0,80	1	0,32	0,00	1,00
25-29	0,35	0,50	1	0,57	-0,50	1,00
25-30	0,58	0,70	1	0,49	-0,50	1,00
29-30	0,35	0,70	1	0,76	-0,80	1,00
30-31	0,50	0,80	0,8	0,57	-0,60	0,80

Elaboración propia

Al analizar las relaciones de causa-efecto se descubrió que por cada dimensión se van relacionando los diferentes factores de dos en dos hasta llevar a uno como efecto final en cada dimensión. Como se muestra en el mapa cognitivo de la dimensión relacionada a la educación (véase la figura 1), en el cual el factor 11, accesibilidad web, es influenciado por todos los factores de esa dimensión en las relaciones de causa-efecto, pero no es la causa de ningún factor.

Se muestra el diagrama de la primera dimensión, la que está relacionada a la educación, para distinguir el grado de complejidad que tiene del proceso expuesto.

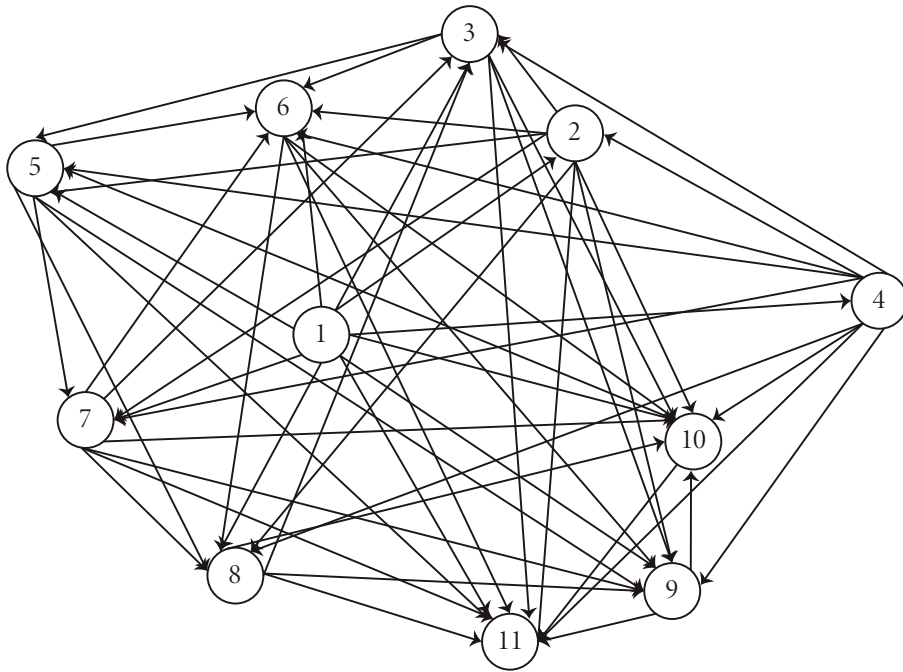


Figura 1. Mapa cognitivo de la dimensión relacionada a la educación  
Elaboración propia

Al observar en la figura 1 y en la tabla 2 que el factor 11 es el resultante de la dimensión relacionada a la educación, se ve que todos los factores dentro de esta dimensión afectan con intensidad, 0,70 o más de mediana, al factor 11, accesibilidad web. Se mencionan al factor 9 y 10 que influyen con mediana 1, el mayor valor, esto es lógico al ser estos factores: accesibilidad *hardware* y accesibilidad *software*; los cuales están relacionados con el factor 11. También se aprecia a los factores 6 y 8 con mediana 0,90, percepción de aprendizaje del estudiante; y recursos formativos y actividades de aprendizaje; los cuales son de gran importancia en la selección de la plataforma virtual y la accesibilidad web que esta pueda tener.

En la dimensión que agrupa los elementos relacionados al sistema, se aprecia como las relaciones de causa-efecto llevan al factor prestigio del *software* como efecto final. De la misma forma, para la dimensión con los elementos relacionados a la organización, se obtiene el factor capacidad de flexibilidad de la organización como efecto final. Estos tres ingredientes se



encuentran relacionados con el componente 31 que consiste en la decisión misma, la selección de la plataforma virtual, siendo los elementos 11, 22 y 29 las causas; y el 31, el efecto. En la segunda ronda también se levantó información referente a la intensidad y relación que indican estos tres factores finales de cada dimensión con respecto a la selección de la plataforma virtual, obteniendo los valores que se muestran en la tabla 2. Utilizando el valor de la mediana como número difuso (Holzmüller y Schlüchter, 2002) y el signo empleado por la mayoría de los expertos, se distingue el mapa cognitivo simplificado de la selección de plataforma virtual educativa en la figura 2.

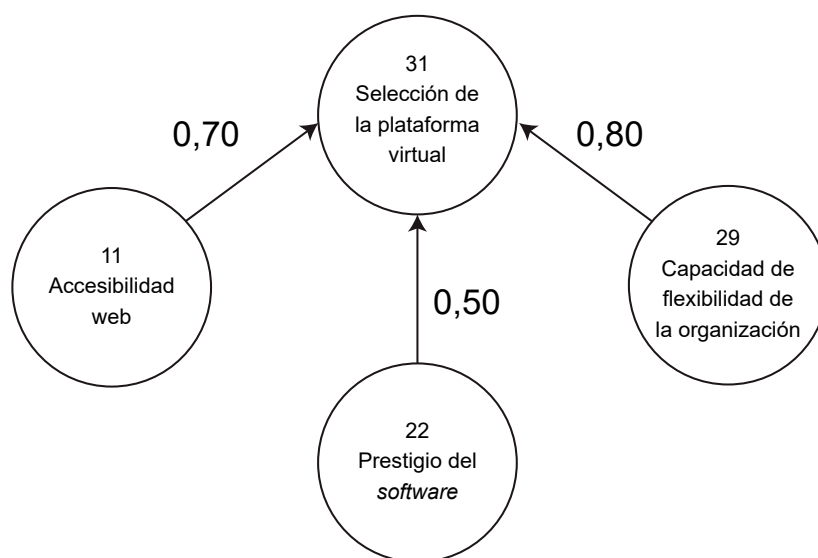


Figura 2. Mapa cognitivo simplificado de la selección de plataforma virtual educativa  
Elaboración propia

En la figura 2 se puede apreciar que el prestigio del *software* influye medianamente en la selección de la plataforma virtual, y que la capacidad de flexibilidad de la organización, así como la accesibilidad web tienen gran influencia en la decisión.

#### 4. CONCLUSIONES

Se aprecia como la selección de la plataforma virtual educativa es influenciada directamente por tres factores, cada uno procedente de una dimensión distinta. El prestigio del *software* es el factor efecto resultante de la dimensión relacionada al sistema, demostrando como todos los elementos

de esta dimensión afectan al prestigio del *software*. De la misma manera para la dimensión con los componentes relacionados a la organización, donde el factor efecto de las conexiones causa-efecto es la capacidad de flexibilidad de la organización. En el caso de la dimensión que agrupa los factores relacionados a la educación, la accesibilidad web, accesibilidad *software* y accesibilidad *hardware* son observados por los expertos como los factores resultantes de esta dimensión, siendo la accesibilidad de *hardware* causa de la accesibilidad *software* y esta de la accesibilidad web, así como de los demás elementos. Es importante mencionar que sin *hardware*, *software* y acceso web no se puede ofrecer un curso virtual, por lo que es razonable descubrir estas relaciones de causa-efecto como las resultantes de la dimensión referida.

También observamos en los datos como el factor de organización y planificación del curso como influencia con gran intensidad a todos los factores, con excepción del *hardware* y *software* que los afecta levemente. Ello se debe a que la planificación del curso determinará la dificultad del mismo, las actividades a realizar que generan la interacción con el profesor, así como las evaluaciones y el diseño didáctico. La percepción de aprendizaje de los estudiantes y la claridad de la comunicación se ve afectada por la organización y planificación del curso, independientemente de los contenidos que este ofrezca.

Se aprecia que la dificultad del curso afecta la claridad de la comunicación y la interacción entre el profesor y el estudiante medianamente. Lo que es comprensible, dado que mientras mayor dificultad tengan los alumnos en comprender el curso, buscarán con mayor frecuencia al profesor incrementando la interacción; de igual forma, el profesor ofrecerá más herramientas de comunicación para asegurar el aprendizaje de los alumnos, siempre que se tengan alumnos responsables e involucrados con el curso. Cabe resaltar, que algunos estudiantes se desaniman con las materias difíciles y no recurren a la ayuda del profesor, por este motivo las relaciones mencionadas tienen una intensidad media.

De acuerdo a los expertos, la interacción entre el profesor y el estudiante son fuertemente influenciadas no solo por la organización y planificación del curso y la claridad en la comunicación; sino por el diseño didáctico-instruccional y los recursos formativos y las actividades. Ello nos demuestra el énfasis en las actividades como estrategias de aprendizaje y herramienta de comunicación, justamente porque es un curso virtual, llevaría a pensar que se pierde el contacto humano y esa confianza o *rapport* que debe existir entre el docente y el alumno. Interacción en la que el estudiante se siente con la suficiente seguridad para confiarle al maestro parte de su vida profesional y pedirle consejos, incluso en los alumnos de pregrado, sobre su vida personal.

Otra relación de causa-efecto, que es importante mencionar, es la influencia del factor de percepción de aprendizaje del estudiante. Vemos que el diseño didáctico-instruccional, la organización y la planificación del curso, la claridad en la comunicación, la interacción entre el profesor y el estudiante, y las evaluaciones generan un fuerte impacto con una intensidad de 0,9, sin embargo, la dificultad del curso tiene un impacto mediado de 0,6. Lo que nos muestra con claridad que tiene mayor importancia, la organización, las actividades y la presencia

(virtual) del profesor, en la percepción de aprendizaje del alumno, que la dificultad del contenido mismo del curso. Esto es importante porque se aprecia como la utilización de estrategias de aprendizaje apoyadas por herramientas de la plataforma virtual para poder estar cerca del alumno son mucho más relevantes que el contenido temático que pueda tener el curso.

Podemos concluir que, para la selección de una herramienta virtual educativa, desde la dimensión relacionada a la educación, se tiene que poner énfasis en la organización, en las estrategias de aprendizaje, en un diseño que sea didáctico y con herramientas que faciliten no solo diferentes actividades destinadas a apoyar lo mencionado, sino que refuercen la comunicación entre el estudiante y el docente.

En trabajos futuros, se pueden realizar simulaciones del mapa cognitivo difuso de selección de una plataforma virtual educativa, y así poder mostrar cómo funciona con métodos dinámicos.

## REFERENCIAS

- Axelrod, R. (1976). *Structure of Decision: The Cognitive Maps of Political Elites*. New Jersey: Princeton University Press.
- Bueno, S., y Salmerón, J. L. (2008). Fuzzy modeling Enterprise Resource Planning tool selection. *Computer Standards & Interfaces* (30), 137-147. doi:10.1016/j.csi.2007.08.001
- Clayton, M. J. (1997). Delphi: A technique to harness expert opinion for critical decision-making tasks in education. *Educational Psychology*, 17(4), 373-387. doi:10.1080/0144341970170401
- Ho, C., Ke, W., y Liu, H. (2015). Choice decision of e-learning system: Implications from construal level theory. *Information & Management*, 52(2), 160-169. doi:10.1016/j.im.2014.07.003
- Holzmüller, H. H., y Schlüchter, J. (2002). Delphi study about the future of B2B marketplaces in Germany. *Electronic Commerce Research and Applications*, 1(1), 2-19. doi:10.1016/S15674223(02)00003-0
- Issroff, K., y Eisenstadt, M. (1997). Evaluating a virtual summer school. *Journal of Computer Assisted Learning*, 13(4), 245-252. doi:10.1046/j.1365-2729.1997.00027.x
- Kosko, B. (1986). Fuzzy cognitive maps. *International Journal on Man-Machine Studies*, 24, 65-75. doi:10.1016/S0020-7373(86)80040-2
- Kosko, B. (1988). Bidirectional associative memories. *IEEE Transactions on Systems, Man, and Cybernetics*, 18(1), 49-60. doi:10.1109/21.87054
- Mirghafoori, S. H., Sharifabadi, A. M., y Takalo, S. K. (2018). Development of Causal Model of Sustainable Hospital Supply Chain Management Using the Intuitionistic Fuzzy

- Cognitive Map (IFCM) Method. *Journal of Industrial Engineering and Management*, 11(3), 588-605. doi:10.3926/jiem.2517
- Pandari, A., y Azar, A. (2017). A fuzzy cognitive mapping model for service supply chains performance. *Measuring Business Excellence*, 21(4), 388-404.
- Papageorgiou, E. I., y Salmerón, J. L. (2013). A review of fuzzy cognitive maps research during the last decade. *IEEE Transactions on Fuzzy Systems*, 21(1), 66-79.
- Ramírez-Fernández, M., y Salmerón Silvera, J. L. (2015). Edutool®: Un instrumento para la evaluación y acreditación de la calidad de los MOOCS. *Educación XXI*, 18(2), 97-123. doi:10.5944/educXXI.13233
- Rovai, A., Ponton, M., Derrick, M., y Davis, J. (2006). Student evaluation of teaching in the virtual and traditional classrooms: A comparative analysis. *Internet and Higher Education*, 9(1), 23-35. doi:10.1016/j.iheduc.2005.11.002
- Salmerón, J. L., y López, C. (2012). Forecasting risk impact on ERP maintenance with augmented fuzzy cognitive maps. *IEEE Transactions on Software Engineering*, 38(2), 439-452. doi:10.1109/TSE.2011.8
- Tarhini, A., Hone, K., y Liu, X. (2014). The effects of individual differences on e-learning users' behaviour in developing countries: A structural equation model. *Computers in Human Behavior*, 41, 153-163. doi:10.1016/j.chb.2014.09.020
- Tornese, F., Verriello, R., Gnoni, M., Mossa, G., y Mummolo, G. (2017). Circular economy strategies for electric and electronic equipment: a fuzzy cognitive map. *Environmental Engineering and Management Journal*, 16(8), 1807-1817. doi:10.30638/emj.2017.197

# Evaluación de un sistema de búsqueda de rutas de evacuación eficientes de un establecimiento usando el algoritmo D estrella (D\*)

Walter Steven Pariona-Sánchez  
20140989@aloe.ulima.edu.pe / Universidad de Lima, Perú

Recepción: 16-6-2019 / Aceptación: 9-7-2019

**RESUMEN.** Los desastres naturales como movimientos telúricos han generado interés en varios autores alrededor del mundo sobre el desarrollo de diferentes soluciones relacionadas a sistemas de evacuación. En esta investigación se expone la importancia de implementar un sistema de evacuación inteligente que reconozca la ruta más corta ante un movimiento telúrico real. De esta manera, la investigación llevó a cabo el proceso de construcción de un simulador para encontrar el camino más corto utilizando el algoritmo D estrella. En esta evaluación se midió el tiempo experimental que le tomó al algoritmo encontrar una ruta de evacuación eficiente bajo diferentes entornos al variar el tamaño del establecimiento, la cantidad de obstáculos iniciales y la cantidad de obstáculos colocados en tiempo real. Los resultados del simulador fueron favorables, puesto que logró identificar una ruta eficiente en 22 milisegundos y un recálculo de ruta en 3 milisegundos, para los casos en que se presenten obstáculos que interfieran con el recorrido inicial.

**PALABRAS CLAVE:** sistemas de evacuación, algoritmos de búsqueda, algoritmo D estrella, caminos más cortos, obstáculos dinámicos

## Evaluation of an Efficient Evacuation Route Search System Within an Establishment Using the D Star Algorithm

**ABSTRACT.** Natural disasters such as earthquakes have aroused great interest among several authors around the world, giving rise to different solutions related to evacuation systems. This research presents the importance of implementing an intelligent evacuation system that recognizes the shortest evacuation route in a real earthquake. Based on this, the research aimed to build a simulator to find the shortest evacuation route using the D star algorithm. In this evaluation, the experimental time taken by the algorithm to find an efficient evacuation route was measured within establishments of different sizes with varying quantities of initial obstacles and varying quantities of obstacles placed in real time. The results of the simulator were favorable since it found an efficient route in 22 milliseconds and a recalculated route in 3 milliseconds for the cases in which obstacles interfered with the initial route.

**KEYWORDS:** evacuation systems, search algorithms, D star algorithm, shortest evacuation routes, dynamic obstacles

## 1. INTRODUCCIÓN

En el Perú, los diferentes tipos de acontecimientos relacionados a un siniestro o a una situación de emergencia pueden ocurrir en cualquier momento debido a que es parte del conglomerado de países que se encuentran en el área del cinturón del fuego la cual causa el 90 % de todos los movimientos telúricos alrededor del mundo.

Adicionalmente, el 60 % del sector construcción en el Perú no ha presentado el plan de evacuación al Instituto Nacional de Defensa Civil (INDECI) (Cornejo, 6 de junio de 2015). Esto evidencia la falta de concientización de las organizaciones y su indiferencia por difundir los conocimientos relacionados a evacuaciones seguras.

Por otro lado, aunque existen establecimientos que cuentan con planes de evacuación regularizados por la institución pertinente, esos planes de escape no necesariamente son efectivos, ya que están compuestos solo de señaléticas las cuales no aseguran un proceso de evacuación eficiente en el momento de un simulacro o en un acontecimiento real.

En consecuencia, al no dejar un escenario de evacuación claro y conocido por todas las personas, estas, en un hipotético caso, irán al primer punto de salida que encuentren (Ying, Zi-min, y Jian, 2017). Además, debe destacarse que es muy probable que un individuo siga a la multitud. Esta probabilidad aumenta cuando una persona no logra identificar la salida y percibe a un flujo de personas dirigiéndose hacia una señal. Esto se debe a que los individuos tienden a creer y seguir otras decisiones de evacuadores (Haghani y Sarvi, 2016).

Debido a lo expuesto, surge el principal problema: los planes o guías de evacuación actuales muestran guías de recorrido y caminos confiables hacia los puntos de salida seguros del establecimiento; no obstante, estos planes no son óptimos al identificar la menor distancia de forma dinámica y no presentan rutas alternas en caso de que aparezcan obstáculos en el recorrido.

La presente investigación propone generar un sistema de evacuación mediante la creación de un simulador que brinde la ruta de escape más eficiente y segura a un área fuera de peligro usando el algoritmo de búsqueda de caminos más cortos, *D estrella* ( $D^*$ ).

## 2. ESTADO DEL ARTE

La implementación de sistemas que ayuden a buscar rutas de evacuación eficientes para los establecimientos ha sido de interés para otros autores alrededor del mundo, es decir, existe una preocupación acerca de qué se está haciendo por evitar daños y pérdidas lamentables, principalmente humanas, frente a acontecimientos como son los movimientos telúricos. Esta inquietud ha generado soluciones para abordar el problema relacionado a los planes de evacuación habituales.

## 2.1 Aplicaciones de sistemas de evacuación en diferentes entornos

### 2.1.1 *Koo, Kim, Kim y Christensen (2013)*

Desde el ataque terrorista del 2001 en Estados Unidos, la evacuación en eventos como ataques terroristas, incendios y sismos es centro de atención de profesionales. Sin embargo, la mayoría de los estudios consideran solo poblaciones homogéneas dejando de lado a personas con discapacidades motrices y sensoriales que necesitan de ayuda para evacuar. Además, estos incluso podrían bloquear la evacuación de otras personas debido a su movilidad reducida y el espacio requerido (Koo, Kim, Kim y Christensen, 2013).

Por ello, se propuso el modelo de simulación BUMPEE. Este tiene la capacidad de manejar distintos tipos de agentes de evacuación virtuales como agentes sin discapacidad, con sillas de ruedas, impedimento visual, entre otros. Como objeto de estudio se usó un edificio de 24 pisos diseñado con ascensores adecuados para evacuaciones (Koo *et al.*, 2013).

Debido a que en evacuaciones reales es posible que un ocupante con silla de ruedas cruce un área con la ayuda de otros individuos, el modelo de simulación implementa una rutina de “sistema de amigos” en el que agentes discapacitados reciben ayuda de otros ocupantes. Además, existe la posibilidad de que los agentes en sillas de ruedas usen los ascensores como medio de evacuación (Koo *et al.*, 2013).

En base a esta propuesta se hicieron experimentos con estrategias controladas basadas en agentes y con uso de ascensores. El primer experimento consideró 60 segundos y agentes discapacitados. El resultado mostró que todos los agentes que no usaban sillas de ruedas evacuaron más rápido respecto a una evacuación simultánea. Para el segundo experimento, los agentes discapacitados usaron ascensores. Así, los agentes terminaron la evacuación más rápida comparado a una evacuación con escaleras (Koo *et al.*, 2013).

### 2.1.2 *Hridi, Das, Anjum y Das, 2016*

Bangladesh se ha visto afectada por desastres naturales como huacicos, lluvias y ciclones causando pérdidas de vidas. Debido a esto, este artículo propone una aplicación que guíe al usuario a un lugar seguro en el menor tiempo posible usando su ubicación en tiempo real y mapeando su comportamiento. Esta propuesta va dirigida para usuarios en y fuera de línea. Los usuarios en línea reciben instrucciones basadas en datos en tiempo real de otros usuarios. Según Hridi, Das, Anjum y Das (2016) el comportamiento de la aplicación para estos varía de acuerdo con la ubicación del usuario (a salvo, fuera o dentro de zona de peligro).

Es en el caso de usuarios que se encuentran fuera de línea en donde se destaca la parte innovadora de la metodología. De acuerdo con Hridi *et al.* (2016), para estos usuarios la aplicación

almacena su patrón de movimiento de los catorce últimos días y almacena información sobre la calidad de las autopistas alrededor del área en la cual se ha movido la persona.

Es así como, basada en datos de los últimos siete días, la aplicación crea un modelo de comportamiento del usuario. Usando ese modelo se predicen las rutas más propensas a ser usadas en un siniestro y se almacenan en un grafo. Así, la ruta de evacuación más segura tendrá un menor valor en el grafo. Por último, la aplicación ejecuta múltiples veces un algoritmo de búsqueda y encuentra todas las rutas de evacuación (Hridi *et al.*, 2016).

Por otra parte, para usuarios sin conexión, el servidor analiza movimientos y hace una suposición de las rutas más propensas a ser usadas durante una evacuación. Así, sugerirá a usuarios en línea rutas alternativas que probablemente no sean utilizadas por usuarios sin conexión y, por lo tanto, permitirá una evacuación más rápida (Hridi *et al.*, 2016).

## 2.2 Investigaciones que emplean algoritmos de caminos más cortos más conocidos

### 2.2.1 Pelechano y Badler (2006)

Pelechano y Badler (2006) mencionan que la evacuación en multitud de edificios es usualmente obstaculizada por personas que no conocen su conectividad interna debido a que los ocupantes suelen usar las salidas familiares ignorando las salidas de emergencia. Además, el aumento del estrés genera una reducción general de la conciencia y un incremento de desorientación.

Para esto se desarrolló MACES (por sus siglas en inglés, *multi-agent communication for evacuation simulation*, es decir, comunicación multiagente para la simulación de evacuación), el cual es un sistema distribuido por *agentes* según su propio comportamiento. Cada agente puede tener una personalidad: líder entrenado, líder no entrenado y seguidor no entrenado. *Uno entrenado* va al camino más corto sin problemas, *un líder* explora el edificio usando un algoritmo de búsqueda por profundidad (DFS), *un seguidor* no sabe qué hacer y solo sigue las decisiones de otras personas (Pelechano y Badler, 2006).

MACES recibe como datos de entrada las características del ambiente como las dimensiones, el número de salidas y de agentes, el plano del edificio, el porcentaje de agentes entrenados. Luego, por cada celda del plano, el algoritmo automáticamente genera el camino más corto a cada salida (Pelechano y Badler, 2006).

En una prueba de comparación de métodos de búsqueda, el resultado del algoritmo DFS fue quince veces más rápido al buscar habitaciones en comparación con un algoritmo de búsqueda aleatoria. En una segunda prueba de comparación entre comunicación y no comunicación, el resultado de la simulación con comunicación entre agentes terminó en la mitad del tiempo que le tomó al caso sin comunicación (Pelechano y Badler, 2006).



## 2.3 Investigaciones que toman en cuenta el comportamiento de los evacuados

### 2.3.1 Iizuka y Iizuka (2015)

Proponen un sistema que funciona en los dispositivos móviles ejecutando cálculos distribuidos con el *framework* DCOP (*distributed constraint optimization problem*) el cual no necesita un servidor central. Este sistema intercambia información acerca de la seguridad, además, planifica, negocia y manifiesta la ruta (Iizuka y Iizuka, 2015).

Para lograr la negociación del tiempo y usar el *framework*, se necesita realizar una formalización. Una ruta de evacuación es considerada un recurso donde cada agente tiene una variable para almacenar la ubicación a la cual debe moverse. El agente también decide la posición en el tiempo  $t + 1$  usando el *framework* en el tiempo  $t$  (Iizuka y Iizuka, 2015).

Con el propósito de obtener resultados se usó un simulador de agentes múltiples. El primer experimento simulaba una evacuación en un campus de una universidad, al ejecutar esa evacuación guiada por el sistema, el tiempo de la evacuación disminuyó en un 10 % respecto a una evacuación sin sistema guiado (Iizuka y Iizuka, 2015).

## 3. ANTECEDENTES

Para brindar la solución planteada ante el problema expuesto es importante tener en cuenta el estado del terreno, pues debe verificarse en tiempo real si el entorno presenta obstáculos como paredes derrumbadas, personas accidentadas, etc. Estos obstáculos se reflejan en un plano virtual como el costo óptimo de un nodo y se tiene que monitorear para saber la variación de su costo en cada tiempo  $t$  y así saber si sigue siendo óptimo.

Es por esto que se presentará la lógica del algoritmo *D estrella* ( $D^*$ ). Este algoritmo es un reensamblaje del algoritmo *A estrella* ( $A^*$ ) para convertirse en un algoritmo dinámico ya que los parámetros del costo del arco pueden cambiar durante el proceso de ejecución.

Al inicio, el agente comienza en un estado y se mueve para alcanzar su ubicación final denotado por  $G$ . Una diferencia a recalcar del algoritmo  $D^*$  con el  $A^*$  es que cada estado, excepto  $G$ , tiene un puntero posterior (*backpointer*) a un siguiente estado  $Y$ ; el costo de atravesar un arco desde  $Y$  a un estado  $X$  es un número positivo dado por la función de costo del arco  $c(X, Y)$ . Si  $Y$  no tiene un arco hacia  $X$ , entonces  $c(X, Y)$  es indefinido (Stentz, 1994).

Cabe mencionar que  $D^*$  también mantiene una lista abierta de estados como el  $A^*$ . En esta se almacenan los cambios en la función de costo del arco y los costos de la ruta. Una característica adicional es que cada estado tiene un *tag* asociado  $t(X)$ , en el que  $t(X) = \text{NUEVO}$  si  $X$  nunca ha estado en la lista abierta,  $t(X) = \text{ABIERTO}$  si  $X$  se encuentra en la lista abierta, y  $t(X) = \text{CERRADO}$  si  $X$  ya no se encuentra en la lista abierta (Stentz, 1994).

Otra característica importante de  $D^*$  es que consta de dos funciones principales: PROCESS-STATE y MODIFY-COST. La primera es usada para calcular los costos de caminos óptimos hacia el nodo final; la segunda se emplea para cambiar la función de costo de arco e ingresar los nuevos estados alterados (ocurrencia de obstáculos) a la lista abierta (Stentz, 1994).

Al iniciar, cada estado  $X$  tiene un  $t(X) = \text{NUEVO}$ , el costo de la ruta es cero y el nodo destino se inserta a la lista abierta. Luego, la función PROCESS-STATE es llamada hasta remover el estado actual de la lista o el valor  $-1$  sea retornado lo cual significa que se encontró la ruta o no existe, respectivamente. Después de esto, el agente sigue la ruta trazada por los *back-pointers* hasta alcanzar el nodo objetivo o descubrir la presencia de un obstáculo. Seguidamente, la función MODIFY-COST corrige la función de costo del arco y actualiza la lista abierta. De esta manera, si  $Y$  es un estado en el que se descubre un error, nuevamente se realiza una llamada a la función PROCESS-STATE. Así, una nueva secuencia  $Y$  se construye y el agente sigue los punteros posteriores hasta encontrar el destino final (Stentz, 1994).

Por último, al comparar el algoritmo  $D^*$  con otros algoritmos de búsqueda, se observó que los resultados son óptimos, ya que este replantea un nuevo camino desde la última posición mientras que los otros replantean el camino como una nueva trayectoria global.

#### 4. METODOLOGÍA

Teniendo en cuenta el objetivo y la literatura revisada, se construyó una simulación para encontrar las rutas de escape más eficientes y seguras en un área fuera de peligro implementando el algoritmo de búsqueda del camino más corto  $D^*$ .

A continuación, en la figura 1 se presenta un esquema equivalente a las etapas que fueron necesarias para la construcción de la solución propuesta.

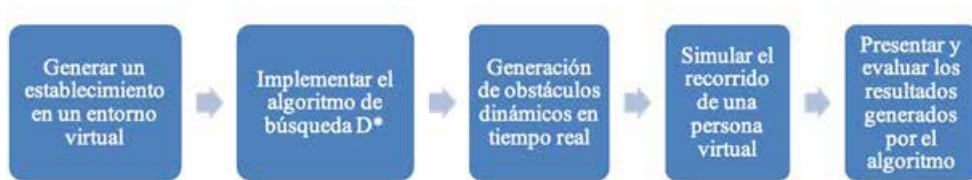


Figura 1. Diagrama de etapas del proyecto de investigación

Elaboración propia

#### 4.1 Proceso de generar un establecimiento en un entorno virtual

Esta es la primera etapa para comenzar a construir la solución propuesta. Se inició con la configuración inicial del entorno para lo cual se utilizó la herramienta multiplataforma Unity3D. Cabe recalcar que el uso de este *software* estuvo en todo el proceso de construir el sistema de evacuación inteligente dado que permite combinar la programación con un entorno visual.

Por otro lado, para generar el plano virtual, se guardaron las dimensiones del plano y la ubicación de cada objeto en una variable de tipo *Vector2*. Además, el sistema dividió el plano en nodos de un metro de diámetro (también llamado, plano en forma de grillas) para representar en cada nodo la ubicación del agente en el espacio.

#### 4.2 Proceso de implementar el algoritmo de búsqueda $D^*$

Esta etapa es la más importante del proceso de construcción del simulador porque involucra la implementación del algoritmo, y su relación y conexión con el plano creado gráficamente en Unity. Para esto se crearon 3 *scripts* con nombre Node, GameGrid y SearchAlgorithm.

El primer *script* sirvió para representar un nodo en memoria e iniciar su posición en el espacio 3D, su posición en el plano, su identificador único, el valor de su función de costo, el valor del *tag*, del puntero posterior y un valor booleano para representar si es un obstáculo.

El segundo *script* manejó la lógica para representar el plano virtual convirtiéndolo en una grilla y encontró los nodos que no son transitables en base a obstáculos iniciales establecidos en la parte gráfica. Aquí se ejecutó el proceso de creación de nodos por cada grilla. Asimismo, una función se encargó de actualizar el plano con el objetivo de almacenar nuevos obstáculos insertados en tiempo de ejecución.

El tercer *script* fue fundamental para la obtención de la ruta óptima ya que manejó la lógica del algoritmo  $D^*$ . Como primer paso se crearon en el motor gráfico dos *GameObject* que representaron la posición del agente y la zona segura más cercana. Luego, se crearon e inicializaron dos arreglos llamados ABIERTO y CERRADO. El primero almacenó los nodos a ser evaluados; el segundo los que ya habían sido evaluados en la ejecución.

Seguido de esto, se definió una función que encontrara la ruta más corta desde la posición del agente hasta la ubicación de la zona de evacuación segura. En esta función el proceso de encontrar la ruta pasa por varias etapas en las que se ejecutan dos funciones fundamentales: *process-state()* y *modify-cost()*. A su vez, también hay funciones auxiliares encargadas de insertar, borrar, obtener el nodo con menor costo y actualizar la búsqueda debido a obstáculos. Cabe mencionar que gracias a las funciones fundamentales, el sistema fue capaz de gestionar ambientes o planos dinámicos ya que es posible insertar obstáculos en tiempo real.

### 4.3 Proceso de generar obstáculos dinámicos en tiempo real

Con esta etapa se implementó una fase exclusiva para la construcción de la lógica que involucra la creación dinámica y aleatoria de obstáculos sobre el plano virtual.

Para esto, además de la implementación del algoritmo de búsqueda  $D^*$ , se programó la función *createRandomObstacles()*. Esta función crea *GameObjects*, define sus posiciones y sus tamaños con base en valores generados aleatoriamente y los establece sobre el plano.

Por último, para que la función *createRandomObstacles()* se pudiera ejecutar continuamente durante la simulación, se hizo uso de la función nativa *InvokeRepeating()* la cual recibió tres parámetros. Como primer parámetro recibió la función *createRandomObstacles()* y la ejecutó cada tres segundos después de cinco segundos de iniciada la simulación.

### 4.4 Simular el recorrido de una persona virtual

En esta etapa el simulador comenzó su proceso de ejecución. Esto debido a que el agente virtual tuvo asignado, como comportamiento o patrón de desplazamiento, el resultado del algoritmo construido y así pudo encontrar rápidamente la ruta más adecuada.

Para entender mejor la ejecución, esta se separó en dos partes. Primero se tuvo que establecer en el plano la posición inicial del agente y la posición del punto de evacuación seguro, para luego ejecutar el simulador y así obtener el recorrido más corto desde la posición inicial hasta la posición final.

Como segunda etapa se insertaron los nuevos obstáculos al plano virtual, en tiempo de ejecución, para conseguir así una nueva ruta en caso interfiriera con la posición actual del agente y la ruta encontrada inicialmente.

A continuación, en las figuras 2 y 3 se muestran los recorridos, tanto el obtenido después de ejecutar el simulador como el replanteado al encontrar un obstáculo, respectivamente.

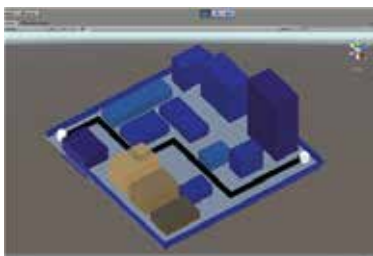


Figura 2. Simulación ejecutada y el camino más corto como resultado  
Elaboración propia

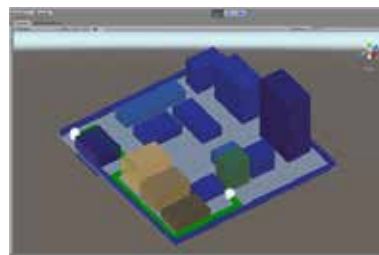


Figura 3. Replanteamiento del camino más corto debido a la presencia de un obstáculo colocado inicialmente  
Elaboración propia

#### 4.5 Presentar y evaluar los resultados generados por el algoritmo

En esta etapa se determinó el tiempo de ejecución experimental que necesitó el sistema para encontrar el camino más corto. Además, se validaron los resultados comparando el tiempo que tomó encontrar el camino al incrementar el tamaño del plano y la cantidad de obstáculos iniciales. Esto con el objetivo de comprobar el correcto funcionamiento y comportamiento del algoritmo.

Por otro lado, también se obtuvo como resultado el nuevo tiempo de ejecución al replantear la ruta debido a la presencia de nuevos obstáculos que interfirieron con el camino encontrado.

Cabe recalcar que la ejecución del simulador se realizó en una computadora con las siguientes características: procesador Intel Core i7-7630QM, CPU 2.60GHz, 8 GB de RAM, sistema operativo MacOS Mojave y tarjeta gráfica Radeon Pro 560X 4096 MB.

### 5. RESULTADOS

Posterior a ejecutar el simulador, en la prueba de concepto se logró calcular el tiempo de ejecución experimental requerido por el algoritmo para encontrar el camino más corto hacia la posición final.

Esto se realizó utilizando un paquete llamado System.Diagnostics el cual “proporciona clases que permiten interactuar con procesos del sistema, registros de eventos y contadores de rendimiento” (Microsoft, s. f.). De esta manera, se logró medir el tiempo transcurrido desde el inicio de la ejecución del simulador hasta la obtención del resultado.

Para la presentación de resultados estos se dividieron en tres criterios. El primer criterio mostró el camino más corto encontrado tomando en cuenta la posición del agente y la posición del punto de extracción seguro. Así, en la figura 4 se puede observar que el CPU requirió 22 milisegundos para encontrar la ruta, es decir, se demoró en calcular el camino más corto 0,022 segundos.

El segundo criterio presentó como resultado el tiempo de ejecución experimental requerido para replantear el nuevo camino más corto ante la existencia de obstáculos agregados manualmente que interfirieron con el camino encontrado inicialmente. En la figura 5 se puede observar que el CPU requirió 2 milisegundos para replantear la ruta más corta ante un obstáculo, es decir, se demoró 0,002 segundos en calcular el nuevo camino.

Por último, el tercer criterio mostró la presencia de obstáculos generados en tiempo real. Así, dio el resultado del tiempo de ejecución experimental requerido para encontrar el nuevo camino más corto ante la existencia de obstáculos generados dinámicamente. En la figura 6 se observa que, con 58 obstáculos dinámicos sobre el plano, el CPU requirió 3 milisegundos o 0,003 segundos para recalcular el nuevo camino más corto.

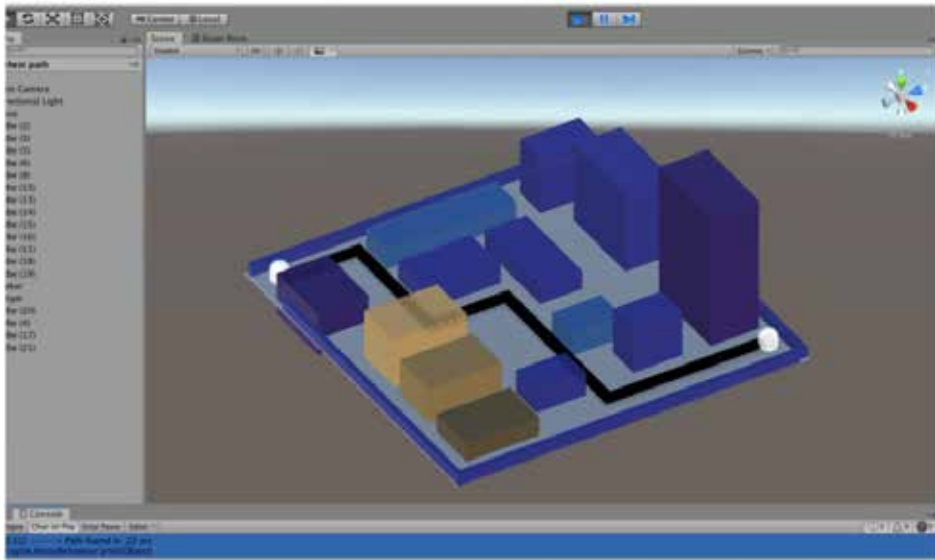


Figura 4. Tiempo requerido en milisegundos para encontrar el camino más corto inicial  
Elaboración propia

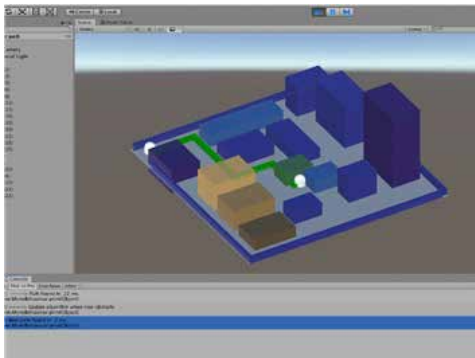


Figura 5. Estado del plano virtual con el nuevo camino más corto recalculado debido a la presencia de un obstáculo  
Elaboración propia

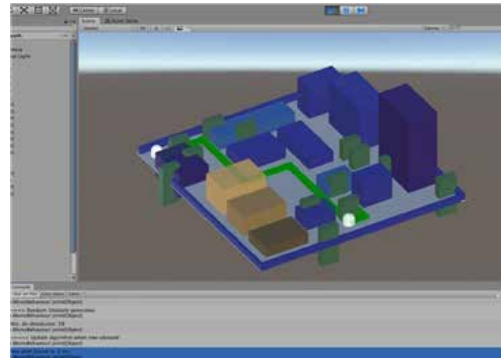


Figura 6. Estado del plano virtual con el nuevo camino más corto recalculado debido a la presencia de obstáculos dinámicos  
Elaboración propia

A manera de resumen, se presenta una tabla comparativa con el objetivo de ilustrar de mejor manera los resultados obtenidos.

Tabla 1

*Tabla comparativa de los resultados obtenidos con relación al tiempo de ejecución*

Algoritmo	Tiempo de ejecución para encontrar el camino más corto inicialmente	Tiempo de ejecución para recalcularse la ruta ante la presencia de un obstáculo agregado manualmente	Tiempo de ejecución para recalcularse la ruta ante la presencia de obstáculos dinámicos
Algoritmo de búsqueda $D^*$	22 milisegundos (ms)	2 milisegundos (ms)	3 milisegundos (ms)

Elaboración propia

## 6. DISCUSIÓN DE LOS RESULTADOS

Para realizar una discusión y validación del funcionamiento del algoritmo y los resultados obtenidos en la prueba de concepto se tuvo en cuenta una técnica recurrentemente utilizada en la literatura revisada con relación a algoritmos más cortos. Esta técnica experimental consiste en realizar diferentes ejecuciones del algoritmo bajo diferentes escenarios.

Para la realización de la validación arbitrariamente se establecieron tres escenarios. El primero con una variación en el tamaño del plano, se realizaron ejecuciones con escalas de 5, 10 y 15 del plano y con 10 obstáculos iniciales.

El segundo escenario con una variación en la cantidad de obstáculos iniciales. Se realizaron ejecuciones con 10, 15 y 20 obstáculos, además, estas simulaciones fueron sobre un plano de escala 5.

Por último, se eligió una combinación de los seis escenarios propuestos. Este escenario ha sido aplicado en otras literaturas para ofrecer conclusiones acerca del algoritmo  $D^*$  ya que se guían del patrón aplicado en Stentz (1994). Así, se generaron cinco ejecuciones sobre un plano de escala 10 con 15 obstáculos iniciales. Con esto se comprobó la consistencia del tiempo de ejecución obtenido en las ejecuciones realizadas.

### 6.1 Primer escenario

A continuación, se presenta una tabla comparativa de los resultados obtenidos al ejecutar el simulador con los planos de tamaño de escala de 5, 10 y 15.

Tabla 2  
*Tabla comparativa de resultados para validar el comportamiento en diferentes dimensiones*

Dimensiones del plano virtual	Tiempo de ejecución experimental obtenido (ms)
Escala 5	101 ms
Escala 10	543 ms
Escala 15	1474 ms

Elaboración propia

Como se puede observar en la tabla 2, el tiempo de ejecución experimental tiene un comportamiento incremental ya que este aumenta conforme se vayan aumentando las escalas del plano virtual.

## 6.2 Segundo escenario

A continuación, la tabla comparativa de los resultados obtenidos al ejecutar el simulador con 10, 15 y 20 obstáculos (véase la tabla 3).

Tabla 3  
*Tabla comparativa de resultados para validar el comportamiento con diferentes obstáculos*

Cantidad de obstáculos sobre el plano virtual	Tiempo de ejecución experimental obtenido (ms)
10	99 ms
15	103 ms
20	104 ms

Elaboración propia

Como se puede observar en la tabla 3, el tiempo de ejecución experimental no varía de manera drástica comparado al escenario anterior. Así, se observa que la cantidad de obstáculos iniciales afectan al tiempo de ejecución en menor proporción comparado a la variación del tamaño del plano. Además, el tiempo de ejecución experimental depende de dónde se ubiquen los obstáculos iniciales en relación al punto de origen (agente) y el nodo objetivo.

## 6.3 Tercer escenario

Como último escenario, se muestra la tabla comparativa que abarca los resultados obtenidos al ejecutar un mismo escenario en cinco iteraciones (véase la tabla 4).



Tabla 4

*Tabla comparativa de resultados para validar la consistencia del algoritmo al obtener el tiempo de ejecución experimental*

Número de ejecución	Tiempo de ejecución experimental obtenido (ms)
1	612 ms
2	607 ms
3	621 ms
4	612 ms
5	605 ms

Elaboración propia

En base a lo observado en la tabla 4, al generar cinco ejecuciones del mismo escenario se obtienen tiempos de ejecución experimentales semejantes cuya variación no excede a 16 ms. De esta manera, se pudo inferir que el algoritmo  $D^*$  ofrece consistencia en la búsqueda del camino más corto. Por otro lado, estas variaciones del tiempo de ejecución pueden ser a causa del desempeño de la CPU en el instante en el que se ejecutó la simulación.

## 7. CONCLUSIONES

Como conclusión, se observa que en los diferentes escenarios realizados en el acápite 5, los resultados indicaron que al usar el algoritmo  $D^*$  se consiguió el camino más corto en solo 22 milisegundos, lo cual indica que en un entorno real de evacuación una persona conocería de manera precisa el camino más corto en menos de un segundo.

Asimismo, también se logró calcular el tiempo experimental que tomaría el algoritmo en recalculer el camino más corto y seguro en caso se evidencie la presencia de obstáculos que interfirieran en el recorrido. Así, al algoritmo solo le tomó 3 milisegundos en recalculer una segunda ruta eficiente. Esto indica que en un entorno real de evacuación una persona, que se encuentre recorriendo el camino sugerido inicialmente, podría conocer en menos de 0,001 segundos el camino más corto y seguro en caso de la presencia de obstáculos.

## REFERENCIAS

Cornejo, M. B. (6 de junio del 2015). 60 % de instituciones sin plan para evacuar. *Correo*. Recuperado de <https://diariocorreo.pe/peru/60-de-instituciones-sin-plan-para-evacuar-592886/>

- Haghani, M., y Sarvi, M. (2016). Human exit choice in crowded built environments: Investigating underlying behavioural differences between normal egress and emergency evacuations. *Fire Safety Journal*, 85, 1-9. doi:10.1016/J.FIRESAF.2016.07.003
- Hridi, A. P., Das, D., Anjum, M. M., y Das, T. (2016). Faster evacuation after disaster: Finding alternative routes using probable human behavior. *ACM DEV'16, Proceedings of the 7th Annual Symposium on Computing for Development*, 1-4. doi:10.1145/3001913.3006632
- Iizuka, Y., y Iizuka, K. (2015). Disaster evacuation assistance system based on multi-agent cooperation. *HICSS'15: Proceedings of the 2015 48th Hawaii International Conference on System Sciences*, 173-181. doi:10.1109/HICSS.2015.30
- Kocay, W., y Kreher, D. L. (s. f.). *Graphs, algorithms, and optimization*. Recuperado de [https://scholar.google.com.pe/scholar?q=Kocay,+W.,+y+Kreher,+D.+L.+\(s.+f.\).+Graphs,+algorithms,+and+optimization.&hl=es&as\\_sdt=0&as\\_vis=1&oi=scholart](https://scholar.google.com.pe/scholar?q=Kocay,+W.,+y+Kreher,+D.+L.+(s.+f.).+Graphs,+algorithms,+and+optimization.&hl=es&as_sdt=0&as_vis=1&oi=scholart)
- Koo, J., Kim, Y. S., Kim, B.-I., y Christensen, K. M. (2013). A comparative study of evacuation strategies for people with disabilities in high-rise building evacuation. *Expert Systems with Applications: An International Journal*, 40(2), 408-417. doi:10.1016/j.eswa.2012.07.017
- Masudur Rahman Al-Arif, S. M., Iftekhharul Ferdous, A. H. M., y Hassan Nijami, S. (2012). Comparative study of different path planning algorithms: A water based rescue system. *International Journal of Computer Applications*, 39(5), 975-8887. Recuperado de <https://pdfs.semanticscholar.org/e59d/c7af0604fced7d124d4f755917725c7892e9.pdf>
- Microsoft. (s. f.). System. Diagnostics Namespace. Recuperado de <https://docs.microsoft.com/en-us/dotnet/api/system.diagnostics?view=netframework-4.8>
- Pelechano, N., y Badler, N. (2006). Modeling Crowd and Trained Leader Behavior during Building Evacuation. *IEEE Computer Graphics and Applications*, 26(6), 80-86. doi:10.1109/MCG.2006.133
- Stentz, A. (1994). Optimal and efficient path planning for partially-known environments. *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA'94)*, 4, 3310-3317. Recuperado de <https://www.ri.cmu.edu/publications/optimal-and-efficient-path-planning-for-partially-known-environments/>
- Thulasiraman, K., y Swamy, M. N. S. (2011). Basic concepts. En Thulasiraman y M. N. S. Swamy, *Graphs: Theory and Algorithms* (pp. 1-30). John Wiley & Sons. doi:10.1002/9781118033104.ch1
- Ying, Z., Zi-min, Z., y Jian, C. (2017). EvacAgent: A Building Emergency Evacuation Simulation Model Based on Agent. *AICT'17: Proceedings of the 2017 International Conference on Artificial Intelligence, Automation and Control Technologies*, 1-7. doi:10.1145/3080845.3080872

# Detección de intrusiones basada en modelado de red resistente a evasión por técnicas de imitación

Jorge Maestre-Vidal  
jmaestre@ucm.es/ Universidad Complutense Madrid, España

Marco Antonio Sotelo-Monge  
masotelo@ucm.es/ Universidad Complutense Madrid, España

Recepción: 11-6-2019 / Aceptación: 9-7-2019

**RESUMEN.** Los sistemas de red emergentes han traído consigo nuevas amenazas que han sofisticado sus modos de operación con el fin de pasar inadvertidos por los sistemas de seguridad, lo que ha motivado el desarrollo de sistemas de detección de intrusiones más eficaces y capaces de reconocer comportamientos anómalos. A pesar de la efectividad de estos sistemas, la investigación en este campo revela la necesidad de su adaptación constante a los cambios del entorno operativo como el principal desafío a afrontar. Esta adaptación supone mayores dificultades analíticas, en particular cuando se hace frente a amenazas de evasión mediante métodos de imitación. Dichas amenazas intentan ocultar las acciones maliciosas bajo un patrón estadístico que simula el uso normal de la red, por lo que adquieren una mayor probabilidad de evadir los sistemas defensivos. Con el fin de contribuir a su mitigación, este artículo presenta una estrategia de detección de intrusos resistente a imitación construida sobre la base de los sensores PAYL. La propuesta se basa en construir modelos de uso de la red y, a partir de ellos, analizar los contenidos binarios de la carga útil en busca de patrones atípicos que puedan evidenciar contenidos maliciosos. A diferencia de las propuestas anteriores, esta investigación supera el tradicional fortalecimiento mediante la aleatorización, aprovechando la similitud de paquetes sospechosos entre modelos legítimos y de evasión previamente construidos. Su eficacia fue evaluada en las muestras de tráfico DARPA'99 y UCM 2011, en los que se comprobó su efectividad para reconocer ataques de evasión por imitación.

**PALABRAS CLAVE:** anomalías, ataques de evasión, detección de intrusiones, redes de comunicación

## Intrusion Detection Based on Evasion-Resistant Network Modeling by Imitation Techniques

**ABSTRACT.** Emerging network systems have brought new threats that have sophisticated their modes of operation in order to go unnoticed by security systems, which has led to the development of more effective intrusion detection systems capable of recognizing anomalous behaviors. Despite the effectiveness of these systems, research in this field reveals the need for their constant adaptation to changes in the operating environment as the main challenge to face. This adaptation involves greater analytical difficulties, particularly when dealing with threats of evasion through imitation methods. These threats try to hide malicious actions under a statistical pattern that simulates the normal use of the network, so they acquire a greater probability of evading defensive systems. In order to contribute to its mitigation, this article presents an imitation-resistant intrusion detection strategy built on the basis of PAYL sensors. The proposal is based on building network usage models and, from them, analyzing the binary contents of the payload in search of atypical patterns that can show malicious content. Unlike previous proposals, this research overcomes the traditional strengthening through randomization, taking advantage of the similarity of suspicious packages to previously constructed legitimate and evasion models. Its effectiveness was evaluated in 1999 DARPA and 2011 UCM traffic samples, in which it was proven effective in recognizing imitation evasion attacks.

**KEYWORDS:** abnormalities, evasion attacks, intrusion detection, communication networks

## 1. INTRODUCCIÓN

Las soluciones iniciales para la detección de intrusos, basadas en el modelado y análisis de entornos de red, aprovecharon originalmente las estrategias de reconocimiento de patrones capaces de descubrir evidencias de ataques previamente conocidos (García-Teodoro, Díaz-Verdejo, Tapiador y Salazar-Hernández, 2015). Pero la rápida proliferación de las tecnologías dio lugar a una cantidad masiva de amenazas antes no vistas, fomentando así el desarrollo de soluciones alternativas capaces de hacer frente a comportamientos maliciosos. Debido a su eficacia en este contexto, el paradigma de detección de intrusos basado en anomalías se ha consolidado como la base de la mayoría de los sistemas de detección de intrusos en la red (NIDS) existentes (Karami, 2018). Este modo de operación típicamente se basa en la construcción de modelos de uso a partir de observaciones legítimas, para luego monitorizar el entorno operativo en busca de discordancias significativas, las cuales son etiquetadas como “sospechosas”. Entre las diferentes publicaciones que han sentado las bases para su desarrollo, nuestra investigación se centra en la detección de intrusiones basada en PAYL (Wang, Cretu y Stolfo, 2005; Wang y Stolfo, 2004). PAYL se basa en el análisis de la carga útil del tráfico en busca de valores estadísticos atípicos dentro de cada contexto de paquetes. A pesar de la evolución de este método, la revisión de la bibliografía revela aun retos a la hora de operar en los escenarios de comunicación (Hadziomanovic, Simionato, Bolzoni, Zamboni y Etalle, 2012; Viswanathan, Tan y Neuman, 2013), como las dificultades al modelar datos extraídos de fuentes heterogéneas, el alto consumo de recursos computacionales, la escasa adaptabilidad a la no estacionariedad (*concept drift*) y la susceptibilidad a los métodos de evasión basados en aprendizaje automático (*machine learning*) (Pastrana, Orfila, Tapiador y Peris-López, 2014), siendo este último el principal objetivo de esta investigación. Con el fin de contribuir a su mitigación, este artículo presenta una nueva estrategia de detección de intrusos resistente a imitación, construida sobre la base de la familia de sensores PAYL, que intenta reforzar los avances del método APAP (Maestre Vidal, Sandoval Orozco y García Villalba, 2017a). De manera similar a sus predecesores, la propuesta construye modelos de uso de la red y, a partir de ellos, analiza los contenidos binarios de la carga útil de tráfico en busca de patrones discordantes que revelen contenidos maliciosos. A diferencia de las soluciones anteriores, esta investigación supera el tradicional fortalecimiento mediante la aleatorización, aprovechando la estimación de similitud de paquetes entre modelos legítimos y de evasión previamente construidos. Se ha llevado a cabo una amplia experimentación que demuestra la efectividad de esta solución en la detección de ataques de ofuscación basados en imitación. Este documento está dividido en cinco secciones, siendo la primera de ellas la presente introducción. En la sección 2 se revisan los trabajos relacionados y se describe el modelo de ataque de evasión considerado en esta propuesta. La sección 3 presenta un nuevo sistema de intrusiones en red (NIDS), descendiente de la familia de sensores PAYL y reforzados contra ataques de evasión. La sección 4 trata sobre la experimentación realizada y discute los resultados obtenidos. Por último, en la sección 5 se resumen las conclusiones y líneas de trabajo futuro.

## 2. ESTADO DEL ARTE

### 2.1 Detección de intrusiones basado en carga útil

El primer sensor de la familia PAYL fue elaborado por Wang y Stolfo en 2004 (Wang y Stolfo, 2004). Según sus autores, el objetivo inicial era detectar la presencia de un gusano (*worm*), ya sea a nivel de puerta de enlace o dentro de una red protegida, evitando así su propagación. Aunque el problema a resolver se basaba en el reconocimiento de gusanos, la propuesta también era válida para una amplia gama de intentos de intrusión, lo que inspiró nuevas líneas de investigación. PAYL se caracterizó por construir un modelo de uso legítimo a partir de 256 características interrelacionadas representadas como histogramas de 256 elementos. Estas se extrajeron según la metodología *n-gram* (Sidorov *et al.*, 2014), como *1-grams*, sobre la distribución de frecuencia de *bytes* presentes en la carga útil. En la etapa de detección, se comparó la similitud entre el modelo normal construido en la etapa de entrenamiento con el modelo generado por el tráfico entrante. Si su divergencia superaba un umbral predefinido, se comunicaba una alerta. Algunos problemas de PAYL inherentes a la construcción de modelos fueron resueltos por una solución impulsada por un mapa autoorganizado (SOM) (Bolzoni Etalle, Hartel y Zambon, 2006). Perdisci, Ariu, Fogla, Giacinto y Lee (2009) propusieron un conjunto de máquinas de vectores de soporte (SVM) para mejorar la precisión del NIDS, siendo abordado de forma similar por modelos ocultos de Markov en Ariu Tronci, y Giacinto (2011). Según esta investigación, las publicaciones anteriores no fueron capaces de reconocer con precisión ataques como *cross site scripting* y *SQL-injection*, donde las estadísticas de carga útil no eran significativamente diferentes del tráfico normal. Este problema fue el principal objeto de otro estudio (Swarnkar y Hubballi, 2016) en donde la efectividad del detector aumentó al implementar la clasificación bayesiana multinomial de una clase. Asimismo, estos sensores introdujeron el muestreo aleatorio con fines de mejora del rendimiento (Ariu *et al.*, 2011; Bolzoni *et al.*, 2006), que se convirtió en una estrategia habitual para reducir el impacto de la inspección profunda de paquetes (del inglés, *deep packet inspection* o DPI) en entornos de comunicación reales.

### 2.2 Detección de *malware* en la carga útil

En respuesta a los retos inherentes a los entornos de monitorización emergentes, el Advanced Payload Analyzer Preprocessor (APAP) (Maestre Vidal *et al.*, 2017a) introdujo una variante más compleja del detector ANAGRAM. APAP se desarrolló originalmente como un módulo de preprocesamiento de Snort, combinando así las capacidades basadas en reglas de dicho NIDS con una nueva capacidad de detección basada en anomalías. Este último se basó en *n-gram* para la extracción de las características de la carga útil (Sidorov *et al.*, 2014). Por otro lado, su almacenamiento/acceso fue gestionado mediante estructuras Counting Bloom Filters (CBF)

(Shana y Venkatachalam, 2014), lo que redujo el consumo de memoria del sistema y mejoró la aplicación de funciones de *hashing*. APAP consta de cinco etapas diferentes de procesamiento de datos, que se agrupan en un par de conjuntos de acciones: entrenamiento y detección. En el entrenamiento se distinguen cuatro fases: inicialización, entrenamiento base, entrenamiento de referencia y definición de valores  $K$ . Durante la inicialización, APAP procede a eliminar información de entrenamientos anteriores, inicializando los CBF y estableciendo la función de *hashing* adecuada. En el entrenamiento base, el CBF se llena con información extraída de muestras de carga útil de tráfico normal. Esta estructura de datos almacena la ocurrencia de cada posible  $n$ -gram en el contenido binario de la carga útil. En el entrenamiento de referencia, se calculan los valores  $K$  del sensor, que son métricas que resumen el contenido de la CBF y facilitan la generación de reglas de detección. Con este fin, se considera un conjunto de datos de muestras “maliciosas”, que sirven como relleno (*padding*) de una réplica del CBF modificado en el entrenamiento base. Los valores  $K$  resultantes se traducen en reglas de detección en la definición de valores  $K$ , que se basa en contrastar los CBF de contenido “normal” y “malicioso”. Finalmente, en el modo detección de APAP, estas reglas determinan la activación de alertas y/o contramedidas.

### 2.3 Evasión basada en imitación

Durante la última década, la comunidad investigadora ha variado su percepción de los ataques basados en imitación del entorno operativo para el que fueron diseñados. En Jonathon, Somesh y Miller (2006) se presentó una de las visiones preliminares de la imitación para evadir los IDS basados en la modelización y el análisis de secuencias de llamadas de sistemas mediante la intercalación de acciones típicamente legítimas entre acciones maliciosas. Investigaciones posteriores (Maestre Vidal, Sandoval Orozco y García Villalba, 2016; Tapiador y Clark, 2010) se enfocaron en cómo fortalecer los sistemas de detección interna contra estas amenazas. Los ataques de imitación también fueron revisados en profundidad en el campo del reconocimiento de intrusiones mediante el análisis del contenido de la carga útil de los paquetes de una red (Pastrana *et al.*, 2014; Wang, Parekh y Stolfo, 2006), lo que supone hoy un reto a considerar. Teniendo en cuenta la representación de ataques de imitación presentada en Jonathon *et al.* (2006), estas amenazas pueden ser entendidas como acciones que intentan explotar la situación descrita en la figura 1; donde  $\Sigma$  es el conjunto de  $n$ -grams extraíbles de una carga útil de paquetes, y  $\Sigma^*$  es el conjunto infinito de todas las cargas útiles posibles. Un modelo legítimo acepta los  $n$ -grams de  $M(L)$  como normales, mientras que los  $n$ -grams de  $M(A)$  son etiquetados como potencialmente dañinos. Por consiguiente, las observaciones que caen en la intersección entre  $M(L)$  y  $M(A)$  conducirán a etiquetados no deterministas. Cuanto mayor es la cercanía a  $M(L)$ , mayor es la probabilidad de ser etiquetado como normal.

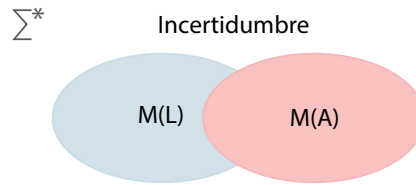


Figura 1. Observaciones potencialmente inadvertidas por el NIDS  
Elaboración propia

En el contexto de PAYL, los ataques de imitación intentan explotar la incertidumbre inherente a la región  $M(L) \cap M(A)$  ocultando el contenido de *malware* dentro de un envoltorio de relleno extraído de  $M(L)$ . Para evitar las técnicas de imitación se asumió que el atacante tiene un conocimiento adecuado de  $M(L)$ , el modelado del uso de la red objetivo y las estrategias de detección implementadas. En respuesta a las capacidades de PAYL, los atacantes adoptaron una amplia variedad de métodos de evasión, siendo la imitación una de sus tácticas más destacadas. Como se demostró en Fogla, Sharif, Perdisci, Kolesnikov y Lee (2006). Frente a ello, la propuesta de ANAGRAM (Wang *et al.*, 2006) introdujo un clasificador binario que implementó filtros de Bloom (Rottenstreich y Keslassy, 2015) para registrar la distribución de los datos de la carga útil, permitiendo así operar sobre *n-grams* de mayor tamaño. ANAGRAM propuso un modelo de *n-gram* aleatorizado para dificultar la generación de muestras de evasión, el cual fue probado como una solución muy precisa, pero la evolución de los métodos de evasión basados en aprendizaje automático mostró que dicho enfoque resultaría insuficiente (Pastrana *et al.*, 2014). Con ello, el fortalecimiento de nuestra propuesta supera la aleatorización al estimar la similitud de paquetes sospechosos entre  $M(L)$  y los modelos de evasión.

### 3. ANÁLISIS DE CARGA ÚTIL DE TRÁFICO PARA LA DETECCIÓN DE INTRUSIONES ROBUSTO A EVASIÓN

Como mejora de los métodos APAP/APACS (Maestre Vidal *et al.*, 2017a, 2017b), el método de fortalecimiento presentado en este trabajo aborda el siguiente modo de operación: en el entrenamiento, tanto los modelos normales como los de evasión se construyen de acuerdo con las características extraídas por la metodología *n-gram* y se almacenan como CBF. En la etapa de detección, las cargas útiles a analizar se recogen del entorno protegido y se comparan con los modelos de uso. Las medidas de similitud entre las observaciones y los modelos de uso de la red previamente construidos en la etapa de entrenamiento permiten estimar su naturaleza (normal o sospechosa) y la coherencia del etiquetado (véase la figura 2). En esta investigación se asume que a mayor diferencia, mayor es la probabilidad de que las muestras fueran elaboradas por los intrusos. Esta sección describe cada etapa del procesamiento de datos y los criterios de decisión adoptados.



### 3.1 Entrenamiento base

En la etapa de entrenamiento se construye el modelo que resume las principales características de las cargas útiles legítimas extraídas de la red. Para su elaboración es necesaria una colección de muestras representativas del tráfico legítimo (habitual) de la red. El modelo de uso normal utiliza CBF para almacenar la frecuencia de ocurrencia de cada  $n$ -gram dentro de la carga útil, alimentando así el CBF(L) con la información extraída de las muestras de referencia hasta que sea posible concluir que su contenido es suficientemente representativo. Se supone que aquello ocurre cuando se añade nueva información y no hay variaciones representativas en la distribución de datos dentro del CBF. Este grado de saturación se evalúa implementando y adaptando el método elbow (Green, Staffell y Vasilakos, 2014), donde el punto de inflexión se calcula observando la suma de errores cuadráticos (SSE) entre las posiciones de los CBF(L) ocupadas por el tráfico normal.

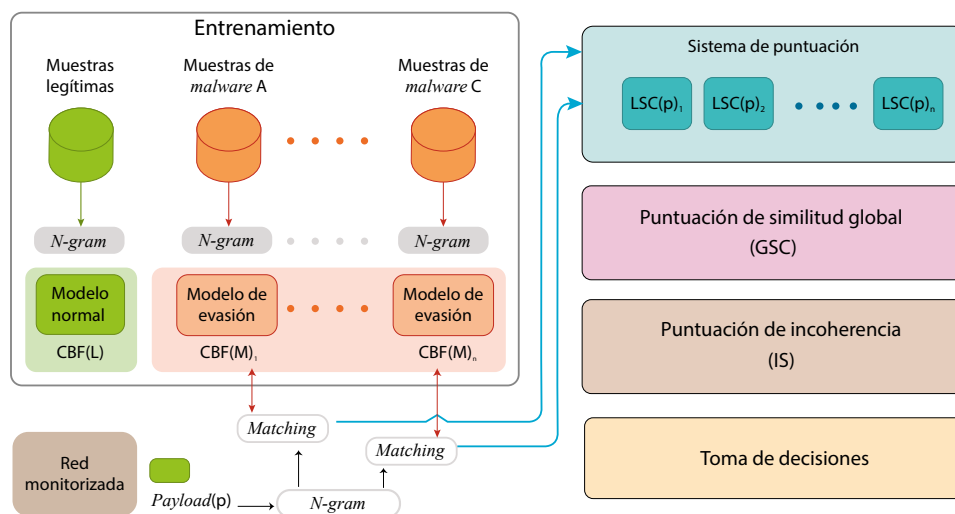


Figura 2. Esquema de procesamiento de datos  
Elaboración propia

### 3.2 Refinamiento por modelos de evasión

Con el fin de reducir la tasa de falsos positivos, en la etapa de entrenamiento refinado, la propuesta genera modelos de evasión capaces de representar los contenidos binarios de diferentes rasgos de *malware*. Siguiendo el mismo procedimiento aplicado para la construcción de modelos de uso normal, se rellena correctamente un CBF por cada grupo de muestras

maliciosas. Sean  $n$  las caracterizaciones de tráfico malicioso tomadas en cuenta, se obtienen los modelos de evasión provisionales correctamente alimentados:

$$CBF'(M) := CBF'(M)_1, \dots, CBF'(M)_n \quad (1)$$

Téngase en cuenta que las trazas maliciosas originales pueden contener elementos de tráfico normal sobre los que se transfiere el *malware*. Para intentar minimizar el número de errores de etiquetado de la carga útil, la operación de sustracción se aplica entre cada uno de ellos y el modelo de uso normal de CBF(L):

$$C(M) := CBF'(M)_1 - CBF(L), \dots, CBF'(M)_n - CBF(L) \quad (2)$$

### 3.3 Sistema de puntuación y detección de intrusiones

Sea la carga útil de tráfico  $p$  y de longitud  $m$  a analizar,  $p$ :  $b_1, b_2, b_3, b_4, b_5, b_6, b_7, \dots, b_m$ , en donde cada  $b_i$ ,  $0 < i \leq m$ , representa un *bit*  $[0,1]$  de su contenido, en la que se asume una ventana deslizante *K-gram* para su extracción, las posiciones del CBF indicadas por las secuencias binarias  $b_1, b_2, \dots, b_{k+1}$ , son accedidas en CBF(L) y cada modelo de evasión  $CB(M)_1, \dots, CBF(M)_n$ . A partir de los resultados obtenidos, se calcula la puntuación del paquete, que tiene en cuenta tanto los aspectos binarios como los espectrales. En el primer caso, se obtiene el número de coincidencias (posiciones CBF superiores a 0) respecto a los modelos normales ( $\gamma$ ) y de evasión ( $\delta_1, \dots, \delta_n$ ). En los enfoques basados en espectro se tienen en cuenta las frecuencias de aparición almacenadas en los CBF, siendo  $\alpha$  respecto a  $\gamma$ , y  $\beta_1, \dots, \beta_n$  respecto a  $\delta_1, \dots, \delta_n$ . Sea modelo legítimo CBF(L) y su correspondiente modelo de evasión CBF(M)<sub>i</sub>,  $0 < i \leq n$ , la puntuación de similitud local (LSC) de la carga útil  $p$  se define mediante la siguiente expresión:

$$LSC(p) = \frac{\alpha - \beta_i}{\alpha + \beta_i + \mu}, \quad LSC \in [-1, 1] \quad (3)$$

donde  $\alpha$  es la suma de las ocurrencias de los *n-grams* extraídos de  $p$  y emparejados exitosamente en CBF(L) (cuando la posición del CBF es mayor que 0),  $\beta_i$  es la suma de ocurrencias de los *n-grams* extraídos de  $p$  emparejados exitosamente en  $CBF(M)_i$ , y  $\mu$  es el número de *n-grams* con valor 0 en los CBF. Por otra parte, la puntuación de similitud global (GSC) de  $p$  se define como el mínimo LSC calculado ( $GSC(p) \in [-1, 1]$ ). En la etapa de detección, el NIDS monitorea la carga útil de los paquetes en busca de patrones de contenido malicioso. En particular, el NIDS calcula las puntuaciones local y global (LSC y GSC, respectivamente) con respecto al modelo de uso normal de la red CBF(L) y los modelos de evasión  $CBF(M)_1, \dots, CBF(M)_n$  previamente construidos en la etapa de entrenamiento. Cuando se emiten alertas ( $p$ )  $< \tau$ , siendo  $\tau \in [-1, 1]$  un intervalo de confianza previamente definido que actúa como parámetro de ajuste del sensor.

### 3.4 Fortalecimiento contra imitación

En esta propuesta, es posible separar el contenido normal de las muestras maliciosas una vez que se realizan las operaciones  $CBF'(M)_i = CBF(M)_i - CBF(L)$ , donde cada patrón común entre  $CBF(L)$  y  $CBF(M)_i$  es reducido o eliminado. Consecuentemente, es altamente improbable que una carga útil de  $p$  a analizar muestre una gran similitud con  $(L)$  y algunos de los  $CBF'(M)_1, \dots, CBF(M)_n$  en la etapa de detección. Se planteó como hipótesis que de esta situación se pueden deducir situaciones en las que el NIDS no etiquetó la muestra con suficiente confiabilidad. En consecuencia, se asumió que cuando el sensor opera sobre modelos de evasión limpios, esto plantea una evidencia potencial de ofuscación por imitación, o un rasgo de entrenamiento deficiente. Por consiguiente, durante la experimentación se considera la ocurrencia de un etiquetado incoherente (IL) cuando se cumple la siguiente expresión:

$$\frac{\alpha}{\alpha + \mu_\alpha} > \phi \quad \text{and} \quad \exists_i, \frac{\beta_i}{\beta_i + \mu_\beta} > \phi \longrightarrow \text{IL} \quad (4)$$

donde  $\alpha$  es la suma de ocurrencias de los  $n$ -grams extraídos de  $p$  coincidentes en  $CBF(L)$  (cuando la posición del CBF es mayor que 0);  $\mu_\alpha$  es el número de  $n$ -grams en  $CBF(L)$  establecidos en 0;  $\beta_i$  es la suma de las ocurrencias de los  $n$ -grams extraídos de  $p$  y emparejados exitosamente en  $CBF(M)_i$ ; y  $\mu$  es el número de  $n$ -grams con valor 0. El intervalo de confianza de incoherencia  $\phi$  actúa como parámetro de ajuste para el reforzamiento del sensor contra imitación. La puntuación de incoherencia (IS) se calcula de la siguiente manera:

$$IS(p) = 1 - \left| \frac{\alpha}{\alpha + \mu_\alpha} - \text{Max} \left\{ \frac{\beta_i}{\beta_i + \mu_i} \right\} \right| \quad (5)$$

## 4. EXPERIMENTOS Y RESULTADOS

### 4.1 Datasets y metodología de evaluación

La propuesta fue evaluada sobre los *datasets* DARPA'99 y UCM 2011. DARPA'99 (Lippmann, Haines, Fried, Korba y Das, 2000) proporciona colecciones de muestras reales y sintéticas monitorizadas en un entorno experimental y plantea un estándar funcional que permite establecer comparativas con trabajos anteriores. Como es habitual en la bibliografía, la validación de la propuesta consideró la segunda versión (Hadziomanovic *et al.*, 2012) entrenada en base a capturas de tráfico recogidas durante siete días y separadas en sesiones etiquetadas como "normales" o "ataques". Nótese que las muestras "normales" se utilizaron para construir el modelo  $CBF(L)$ , y los segundos para construir los modelos de evasión  $CBF(M)_1, \dots, CBF(M)_m$ . Con el fin de evaluar las mejoras de la propuesta respecto a APAP y las contribuciones a la familia PAYL, se consideró el conjunto de datos UCM 2011. Este recoge trazas reales de

tráfico monitorizadas en la red de la Facultad de Informática de la Universidad Complutense de Madrid. El registro de tráfico se realizó a lo largo del año 2011 en diferentes intervalos de tiempo y meses. Las muestras de tráfico normal incluyen actividades de uso habitual de la red, entre ellas: intercambios P2P, transferencias de archivos de varios formatos (.doc, .pdf, .mp3, .jpg, etc.) vía SMTP, navegación HTTP/HTTPS, etc. Los contenidos maliciosos se clasificaron en dieciséis grupos diferentes de amenazas, incluyendo varias familias de *malware*, ataques DoS o procedimientos de obtención de privilegios. La experimentación utilizó el generador de ataques de imitación descrito en Maestre Vidal *et al.* (2016). Con el propósito de hacer que la carga útil del tráfico malicioso incremente su similitud con las características del tráfico normal, la herramienta de ofuscación se basó en el CBF(L) construido en la etapa de entrenamiento. La selección de las secuencias de relleno consideró la estrategia de muestreo estocástico *probability proportional-to-size* (PPS).

#### 4.2 Efectividad con DARPA'99

Tal y como se señala en Wang *et al.* (2006) y se ratifica en Hadziosmanovic *et al.* (2012) y Maestre Vidal *et al.* (2017a, 2017b), dada la naturaleza del conjunto de datos DARPA'99, una ventana deslizante de 3 *grams* resultó ser la configuración más adecuada para los sensores inspirados en PAYL, por lo que se mantuvo dicha configuración en esta propuesta. En esta experimentación, el parámetro de ajuste del sensor fue el intervalo de confianza  $\tau \in [-1, 1]$ , por lo que cada carga útil  $p$  se etiquetó como anómala cuando la expresión  $(p) < \tau$  fue satisfecha. Tal y como se muestra en las comparativas anteriores, los resultados obtenidos se evaluaron en función del mejor ajuste entre la tasa de verdaderos positivos (TPR) y la tasa de falsos positivos (FPR) estimada por el índice óptimo de Youden (J). La tabla 1 muestra la eficacia de las publicaciones relacionadas y resume los resultados obtenidos en esta propuesta. Como se indica, la calibración óptima resultó en una tasa de aciertos del 100 % y FPR = 0,01 % cuando  $J = 0,9998$ . Estos resultados son cercanos a los del mejor sensor cuando opera en circunstancias similares (ANAGRAM). Pero a pesar de estos hallazgos, el uso funcional de DARPA'99 carece de una manera de demostrar la robustez de los sensores frente a amenazas de evasión, por lo cual fue necesario realizar experimentos adicionales para evaluar el resto de las mejoras de esta propuesta. Con el fin de evaluar el fortalecimiento de la propuesta frente a las amenazas de evasión, esta prueba consideró un intervalo de confianza fijo de  $\tau$ , en particular, el que mostraba el equilibrio óptimo entre el TPR y el FPR en los experimentos anteriores. El parámetro de calibración fue la variación del índice de incoherencia en el etiquetado  $\phi$ , hipotetizando que cuanto mayor sea  $\phi$ , mejor será la TPR. El experimento suponía que cada incoherencia en el etiquetado ocultaba una carga útil maliciosamente ofuscada.

Tabla 1  
Comparación con propuestas anteriores (DARPA'99)

Propuesta	Porcentaje FPR	Porcentaje TPR
PAYL (Wang y Stolfo, 2004)	0,00	90,76
Poseidon (Bolzoni <i>et al.</i> , 2006)	0,00	92,00
AnPDPP (Thorat Khandelwal, Bruhadeshwar y Kishore , 2009)	0,06	100,00
Anagram (Wang <i>et al.</i> , 2006)	0,00	100,00
McPAD (Perdisci <i>et al.</i> , 2009)	0,33	87,80
RePIDS (Jamdagni Tan, He, Nanda y Liu, 2013)	0,67	99,33
APAP (Maestre Vidal <i>et al.</i> , 2017a)	0,15	100,00
Esta propuesta	0,01	100,00

Elaboración propia

En la figura 3 se ilustra la variación del mejor ajuste de  $\phi$  según el índice de Youden, donde se observa que, tanto en los modos de funcionamiento convencionales como en los reforzados, la propuesta presenta sensibilidad a las variaciones en la longitud del relleno. Sin embargo, el enfoque reforzado se redujo significativamente en el impacto, siendo mínima la precisión observada de aproximadamente  $AUC \approx 0,95$ .

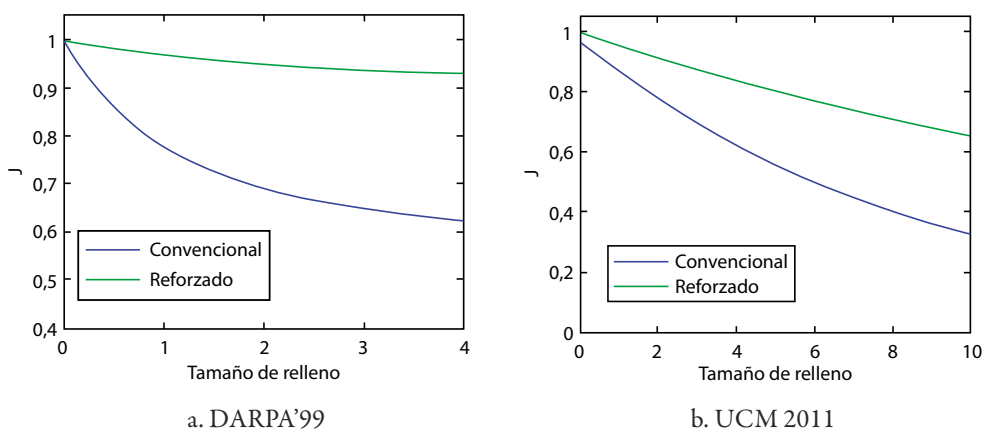


Figura 3. Impacto del tamaño de relleno  
Elaboración propia

### 4.3 Efectividad con UCM 2011

Como se muestra en la tabla 2, la mejor configuración ( $AUC = 0,9900$ ,  $J = 0,9574$ ) superó al mejor ajuste APAP ( $AUC = 0,9136$ ,  $J = 0,8442$ ), pero fue similar a APACS ( $AUC = 0,9902$ ,  $J = 0,9339$ ). Sin embargo, dado que esta propuesta no se basa en correlación de las alertas (como es el caso de APACS), supone una solución más eficaz. Por otro lado, APACS requiere una base de conocimiento correctamente etiquetada, mientras que la presente propuesta proporciona tolerancia a los errores de etiquetado en los datos de entrenamiento. El reforzamiento propuesto contra los intentos de evasión basados en imitación volvió a demostrar su eficacia al aplicarse sobre las muestras de tráfico UCM 2011. Esto se ilustra en la figura 3b, donde se reflejan las variaciones del índice de Youden basadas en la distribución del contenido de relleno. En este experimento se fijó el intervalo de confianza  $\tau$  al ajuste óptimo de la prueba anterior, implementándose una ventana deslizante de 5 *grams*, y el parámetro de sensibilidad fue el índice de incoherencia en el etiquetado  $\phi$ . La versión reforzada no mitigó por completo el impacto de las amenazas de imitación, pero la redujo considerablemente, registrando las peores AUC con valores cercanos a 0,9, lo que mejoró representativamente el 0,7 alcanzado por un despliegue no robusto. El punto de saturación fue de aproximadamente  $AUC \approx 0,85$ , mientras que en la versión original era cercana a 0,6.

Tabla 2  
Resumen de resultados UCM 2011

NIDS	FPR	TPR	AUC	J
Esta propuesta	0,021	0,967	0,9900	0,9574
APAP (Maestre Vidal <i>et al.</i> , 2017a)	0,080	0,947	0,9136	0,8442
APACS (Maestre Vidal <i>et al.</i> , 2017b)	0,034	0,995	0,9902	0,9339

Elaboración propia

## 5. CONCLUSIONES

Este artículo presentó un nuevo enfoque para la detección estadística de *malware* en entornos de comunicación emergentes. La propuesta se basó en el análisis de la carga útil de tráfico en busca de discordancias con respecto a modelos legítimos construidos previamente en la etapa de entrenamiento. La propuesta adoptó las bases de la familia de sensores PAYL, y amplió las soluciones APAP y APACS, aprovechando así la metodología *n-gram* y las estructuras de datos CBF. A diferencia de las soluciones anteriores, esta investigación superó el tradicional fortalecimiento mediante la aleatorización, valiéndose de la similitud de paquetes sospechosos entre modelos legítimos y de evasión construidos previamente. La propuesta ha sido evaluada con las muestras de tráfico DARPA'99 y UCM 2011, demostrando alta precisión y similitud

con las mejores propuestas. Sin embargo, su eficacia demostró ser superior a la de sus antecesores en el procesamiento de carga útil maliciosa ofuscadas por los métodos de imitación. Pero a pesar de los detalles provistos en este artículo, la discusión de algunos otros aspectos fue pospuesta para el trabajo futuro, así como la evaluación de la propuesta sobre entornos de monitorización alternativos.

## REFERENCIAS

- Ariu, D., Tronci, R., y Giacinto, G. (2011). HMMPayL: An intrusion detection system based on hidden Markov models. *Computers y Security*, 30(4), 221-241.
- Bolzoni, D., Etalle, S., Hartel, P., y Zambon, E. (2006). Poseidon: a 2-tier anomaly-based network intrusion detection system. *Proceedings of the 4th IEEE International Workshop on Information Assurance (IWIA)*, 144-156.
- Fogla, P., Sharif, M., Perdisci, R., Kolesnikov, O., y Lee, W. (2006). Polymorphic blending attacks. *Proceedings of the 15th USENIX Security Symposium*, 241-256.
- García-Teodoro, P., Díaz-Verdejo, J. E., Tapiador, J. E., y Salazar-Hernández, R. (2015). Automatic generation of HTTP intrusion signatures by selective identification of anomalies. *Computers and Security*, 55, 159-174.
- Green, R., Staffell, I., y Vasilakos, N. (2014). Divide and Conquer? k-Means Clustering of Demand Data Allows Rapid and Accurate Simulations of the British Electricity System. *IEEE Transactions on Engineering Management*, 61(2), 251-260.
- Hadziosmanovic, D., Simionato, L., Bolzoni, D., Zambon, E., y Etalle, S. (2012). N-gram against the machine: On the feasibility of the n-gram network analysis for binary protocols. *Proceedings of the 15th International Symposium on Recent Advances in Intrusion Detection (RAID)*, 59-81.
- Jamdagni, A., Tan, Z., He, X., Nanda, P., y Liu, R. P. (2013). RePIDS: A multi tier realtime payload-based intrusion detection system. *Computer Networks*, 57, 511-824.
- Jonathon, T., Somesh, J., y Miller, B. P. (2006). Automated discovery of mimicry attacks. *Proceedings of the 9th International Symposium on Recent Advances in Intrusion Detection (RAID)*, 41-60.
- Karami, A. (2018). An anomaly-based intrusion detection system in presence of benign outliers with visualization capabilities. *Expert Systems with Applications*, 108, 36-60.
- Lippmann, R., Haines, J. W., Fried, D. J., Korba, J., y Das, K. (2000). The 1999 DARPA off-line intrusion detection evaluation. *Computer Networks*, 34(4), 579-595.

- Maestre Vidal, J., Sandoval Orozco, A. L., y García Villalba, L. J. (2016). Online masquerade detection resistant to mimicry. *Expert Systems with Applications: An International Journal*, 61, 162-180.
- Maestre Vidal, J., Sandoval Orozco, A. L., y García Villalba, L. J. (2017a). Advanced payload analyzer preprocessor. *Future Generation Computer Systems*, 76, 474-485.
- Maestre Vidal, J., Sandoval Orozco, A. L., y García Villalba, L. J. (2017b). Alert correlation framework for malware detection by anomaly-based packet payload analysis. *Journal of Network and Computer Applications*, 97, 11-22.
- Pastrana, S., Orfila, A., Tapiador, J. E., y Peris-López, P. (2014). Randomized anagram revisited. *Journal of Network and Computer Applications*, 21, 182-186.
- Perdisci, R., Ariu, D., Fogla, P., Giacinto, G., y Lee, W. (2009). McPAD: A multiple classifier system for accurate payload-based anomaly detection. *Computer Networks*, 53(6), 864-881.
- Rottenstreich, O., y Keslassy, I. (2015). The Bloom paradox: when not to use a Bloom filter. *IEEE/ACM Transactions on Networking*, 23(3), 703-716.
- Shana, J., y Venkatachalam, T. (2014). An Improved Method for Counting Frequent Itemsets Using Bloom Filter. *Procedia Computer Science*, 47, 84-91.
- Sidorov, G., Castillo, F., Stamatatos, E., Gelbukh, A., y Chanona-Hernández, L. (2014). Syntactic N-grams as machine learning features for natural language processing. *Expert Systems with Applications: An International Journal*, 41, 853-860.
- Swarnkar, M., y Hubballi, N. (2016). OCPAD: One class Naive Bayes classifier for payload based anomaly detection. *Expert Systems with Applications: An International Journal*, 64, 330-339.
- Tapiador, J. E., y Clark, J. A. (2010). Information-theoretic detection of masquerade mimicry attacks. *2010 Fourth International Conference on Network and System Security*, 183-190.
- Thorat, S. A., Khandelwal, A. K., Bruhadeshwar, B., y Kishore, K. (2009). Anomalous packet detection using partitioned payload. *Journal of Information Assurance and Security*, 3(3), 195-220.
- Viswanathan, A., Tan, K., y Neuman, C. (2013). Deconstructing the Assessment of Anomaly-based Intrusion Detectors. *Proceedings of the 16th International Symposium on Recent Advances in Intrusion Detection (RAID)*, 286-306.
- Wang, K., Cretu, G., y Stolfo, S. J. (2005). Anomalous Payload-based Worm Detection and Signature Generation. *Proceedings of the 8th International Symposium on Recent Advances in Intrusion Detection (RAID)*, 227-246.



- Wang, K., Parekh, J. J., y Stolfo, S. J. (2006). Anagram: A Content Anomaly Detector Resistant to Mimicry Attack. *Proceedings of the 9th International Symposium on Recent Advances in Intrusion Detection (RAID)*, 226-248.
- Wang, K., y Stolfo, S. J. (2004). Anomalous Payload-based Network Intrusion Detection. *Proceedings of the 7th International Symposium on Recent Advances in Intrusion Detection (RAID)*, 203-222.



# La estructura organizacional, el *agile mindset* y el gobierno de TI para la transformación digital de las empresas

Marco A. Fernández del Pomar  
marco9129@gmail.com / Alicorp, Perú

Recepción: 17-6-2019 / Aceptación: 10-7-2019

RESUMEN. El presente artículo describe los bloques fundamentales que las organizaciones deberían considerar para poder realizar, de forma exitosa, una transformación digital. Se describirá como una estructura organizacional adecuada, un cambio cultural hacia el *agile mindset* y una correcta implementación de un gobierno corporativo de TI pueden impulsar a las empresas a alcanzar los cinco dominios de la transformación digital y generar el cambio deseado.

PALABRAS CLAVE: *agile*, agilidad, *agile mindset*, gobierno, gobierno de TI, gobierno corporativo, transformación digital, liderazgo, estructura organizacional, disrupción

## The Organizational Structure, the Agile Mindset and the IT Governance for the Digital Transformation of Companies

ABSTRACT. This article describes the fundamental blocks that organizations should consider in order to successfully carry out a digital transformation. It will be described as an appropriate organizational structure, a cultural shift towards agile mindset, and a correct implementation of a corporate IT governance that can drive companies to reach the five domains of digital transformation and generate the desired change.

KEYWORDS: agile, agility, agile mindset, government, IT governance, corporate governance, digital transformation, leadership, organizational structure, disruption

## 1. INTRODUCCIÓN

Los primeros indicios de lo que hoy conocemos como transformación digital (en adelante, TD) tuvo sus inicios a mitad de los años noventa con el nacimiento del internet. Posteriormente, a mitad de la década pasada, este término comenzó a ganar notoriedad nuevamente. Sin embargo, es recién a principios de esta década que el término comienza a tomar la forma con la que hoy lo conocemos, y es ahí cuando la famosa “ola” de la TD empieza a crecer y a llevarse muchas organizaciones por su paso.

Para aquellas organizaciones que no nacieron como nativas digitales, el subirse a la ola de la TD es una obligación en caso quisieran sobrevivir en la actual cuarta revolución industrial. Esta ola llegó hace varios años para algunas industrias como la banca, telecomunicaciones y tecnologías de la información. Para otras, como la industria de *retail* o consumo masivo, demoró un poco más. Esta tendencia es, sin embargo, algo inevitable y es una transformación por la que todas las empresas en todas las industrias deben pasar. Se trata de un cambio que es muchas veces llamado altamente disruptivo, situación que se vuelve una realidad para aquellas empresas que no son nativas digitales.

A pesar de que no existe una única fórmula para alcanzar la TD con éxito, sí existe actualmente literatura en abundancia sobre el tema (Siebel, 2019; Westerman, Bonnet y McAfee, 2014; Herbert, 2017). Algunas de estas fuentes fueron utilizadas para el presente estudio. Sin embargo, y a pesar de la abundante información al respecto, es aún difícil el identificar aquellas variables que dificultan que una organización “surfee” con éxito la ola de la TD. Este artículo pretende proponer un camino hacia esta transformación basada en tres bloques fundamentales los cuales pueden parecer obvios para algunos, pero que si no son alineados correctamente, a pesar de poseerlos, no servirán de nada para alcanzar el objetivo transformacional si es que no están correctamente alineados.

El camino hacia esta transformación está conformado por la creación de una estructura organizacional como los cimientos, la generación de una cultura ágil, con el *agile mindset* como *core* de la cultura organizacional y, por último, la implementación de un gobierno corporativo de tecnologías de la información (en adelante, GEIT, por sus siglas en inglés) para el alineamiento de los objetivos estratégicos organizacionales con las iniciativas tecnológicas.

Estos tres bloques, si bien son fundamentales para alcanzar una TD exitosa, no cierran el círculo por completo. La etapa final es el alineamiento de estos tres bloques con los cinco dominios de la TD (Rogers, 2016) que se describirán más a detalle en el presente artículo. Es este alineamiento el que permitirá finalmente a las organizaciones alcanzar esta ola, surfearla con éxito y permitir cosechar los beneficios de implementar correctamente aquello que resulta esquivo para muchas organizaciones: una transformación profunda en toda la organización.

## 2. METODOLOGÍA

Desde que se inició la ola de la TD, como la conocemos hoy, a la fecha no existe una definición universalmente aceptada sobre lo que significa el término. Sin embargo, entre muchas de las definiciones existentes que son generalmente aceptadas, existen características muy similares. Una definición propuesta sobre lo que es la TD y que incluye estas características comúnmente aceptadas es la siguiente:

La transformación digital es la integración de tecnología digital en todas las áreas de un negocio, fundamentalmente, cambiando como opera y entregando valor a sus clientes. Es también un cambio cultural que requiere que las organizaciones reten continuamente el *status quo*, experimenten y acepten el fracaso. No se trata de tecnología, se trata de una estrategia aplicada a toda la organización.

Esta definición, y otras existentes en las diversas fuentes que buscan aclarar el término y facilitar a los profesionales el aplicar en la realidad una TD de forma exitosa, son en general bastante fáciles de comprender, pero en la gran mayoría de casos, difíciles de aplicar.

La tecnología es un facilitador que se integra a todo nivel organizacional, pero no es el fin en sí de la transformación. Incluye un cambio cultural que genere valor tanto a los clientes como a la misma organización. Es flexible y acepta el cambio. Se trata de aplicar una estrategia más que implementar un proceso nuevo. Todos los puntos mencionados no son ninguna novedad y se mencionan como conceptos que deben ser aplicados para alcanzar esta transformación. Sin embargo, el aplicarlo suele ser difícil, pues para poder realmente comenzar a ejecutar una estrategia al nivel que requiere la TD, es necesario que previamente existan ciertas condiciones que puedan sustentar estos cambios que, usualmente, suelen ser abruptos o disruptivos para muchas organizaciones que no nacieron en la era digital.

### 2.1 La pirámide

El presente estudio propone enfocarse en tres aspectos organizacionales que abarcan virtualmente a toda la organización y funcionan como base para el tipo de cambio que implica una TD. Estos son la estructura organizacional, el *agile mindset* y el gobierno corporativo de tecnologías de la información (TI). Las sinergias generadas entre estos permitirían a una organización alcanzar un nivel de madurez suficiente para comenzar a generar el cambio necesario que propone una TD. En la figura 1 se muestra una representación de como se unen estos tres bloques a nivel organizacional y como a través de sus sinergias y del alineamiento con los cinco dominios de la TD (Rogers, 2016), se puede finalmente alcanzar la cúspide de la pirámide, la ola de la TD.

El primer bloque, la estructura organizacional, funciona como los cimientos que sostienen el resto de la estructura, que están formados por los bloques del *agile mindset*, el gobierno corporativo de TI, los cinco dominios de la TD (Rogers, 2016) y la cúspide.

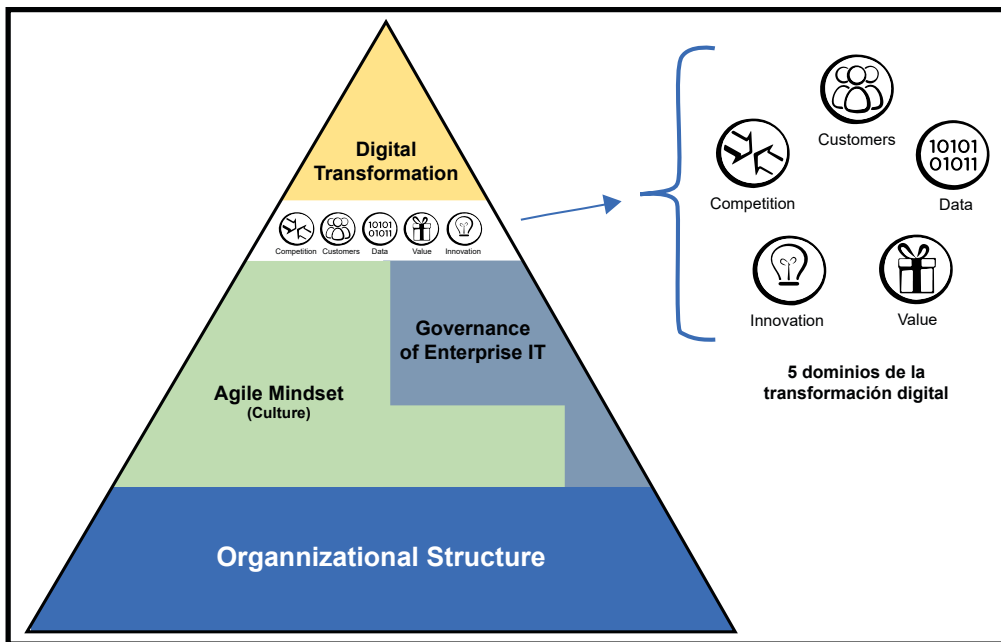


Figura 1. Triángulo de los bloques y dominios para la TD  
Elaboración propia

Lo que usualmente se estila en muchas organizaciones que buscan alcanzar la TD es pretender empezar con un cambio cultural. Esta aproximación al objetivo de la TD tiene un punto válido, el cual es que los colaboradores de todos los niveles organizacionales deben estar alineados, motivados y dispuestos a impulsar la iniciativa para poder generar el cambio cultural en toda la organización. Sin embargo, este inicio flaquea desde dos aristas críticas que pueden generar un revés para la compañía.

El empezar con un cambio cultural, implica usualmente una transformación mental de toda la organización, que es un cambio bastante disruptivo de por sí. Este tipo de cambio suele generar una problemática común que llamaremos *primera arista*; la resistencia al cambio por parte del personal. Esto sucede ya que las personas deben enfrentarse a una incertidumbre que no es del todo tangible para ellas, pero que saben que existe y traerá consigo consecuencias que aun no comprenden completamente. Adicional a la resistencia al cambio por parte del personal, la *segunda arista* que dificulta el generar el cambio cultural, es la falta de una estructura organizacional que lo soporte. Un cambio cultural, especialmente uno que busca generar una TD, implica cambios de pensamiento que, de no ser soportados por un organigrama correctamente alineado al cambio, difícilmente alcanzarán un nivel de madurez suficiente para empezar a cosechar los beneficios de este nuevo *mindset*.

## 2.2 Los CoE aislados, limitados y balanceados

Un ejemplo tangible que se repite con frecuencia en el Perú y en otros lugares es generar centros de excelencia (en adelante, CoE, por sus siglas en inglés) que funcionan como organizaciones independientes y completamente autónomas a su organización matriz. Es decir, estos CoE tienen su propia cultura, políticas y estructura organizacional, que muchas veces difiere totalmente de aquellas de su organización matriz. Si bien un CoE podría permitir generar de forma más rápida iniciativas de TD, al momento de intentar implementar estas soluciones digitales en la organización matriz, que es finalmente el beneficiario y *stakeholder* principal, surgen grandes dificultades inherentes al aislamiento que se genera por tener una estructura organizacional independiente. Esto se debe a que el tratar que una solución digital, creada de forma “externa” y por personal con una cultura organizacional diferente, encaje y se implemente de forma eficiente en una organización que no comparte esta misma cultura, tiene altas probabilidades de fracaso debido a que ambas culturas, procesos y estructuras organizacionales no encajan. El personal que crea dicha solución desde el CoE se encuentra lejos y aislado de la gente del negocio, quienes están en el lado de la organización matriz. Esto se puede considerar como un extremo no deseado, el aislamiento.

El otro extremo que también debe ser evitado es el pretender generar un CoE que solo funcione para una de las áreas de negocio. Por ejemplo, el crear un área de innovación que solo se encargue de iniciativas digitales para el equipo de, por ejemplo, cadena de suministro de la organización. Incluso con una estructura organizacional adecuada, y generando sinergias correctas con el lado de negocio del área, el alcance y, por tanto, el impacto de un CoE de este tipo se vería bastante limitado para poder generar un cambio que conlleve a una TD. Este extremo conlleva a un CoE con grandes limitaciones e incapaz de generar el cambio deseado.

El mejor aproximamiento debería ser el generar un CoE dentro de la misma organización matriz, pero con una estructura organizacional propia de ser necesario, y que no se encuentre limitada a beneficiar a una sola área del negocio, sino que busque impactar a toda la organización. El tener un CoE interno facilita que los dueños finales (*stakeholders*) de las iniciativas digitales que se vayan a crear, sean las propias áreas de negocio. Esto permitirá asegurar un mayor involucramiento y participación de ellos en cada iniciativa digital que se genere en el CoE. Es decir, hacer que el negocio sea 100 % dueño de las iniciativas, pero que estas sean ejecutadas por el CoE. Adicionalmente, y de ser posible, este CoE interno debería tener la suficiente libertad para poder crear sus propios procesos, dinámica de trabajo, e incluso, hasta una subcultura dentro de la organización.

El balance alcanzado por un CoE lo suficientemente independiente, pero que se encuentre dentro de la organización y cerca a las áreas de negocio permitiría generar un alto impacto y beneficios para todas estas. Para poder lograrlo es necesario que se genere una estructura organizacional adecuada para este tipo de CoE, la cual serviría luego como gatillador para el cambio de la estructura organizacional a nivel de toda la institución. La figura 2 muestra los tres ejemplos de CoE que pueden generarse en una organización.

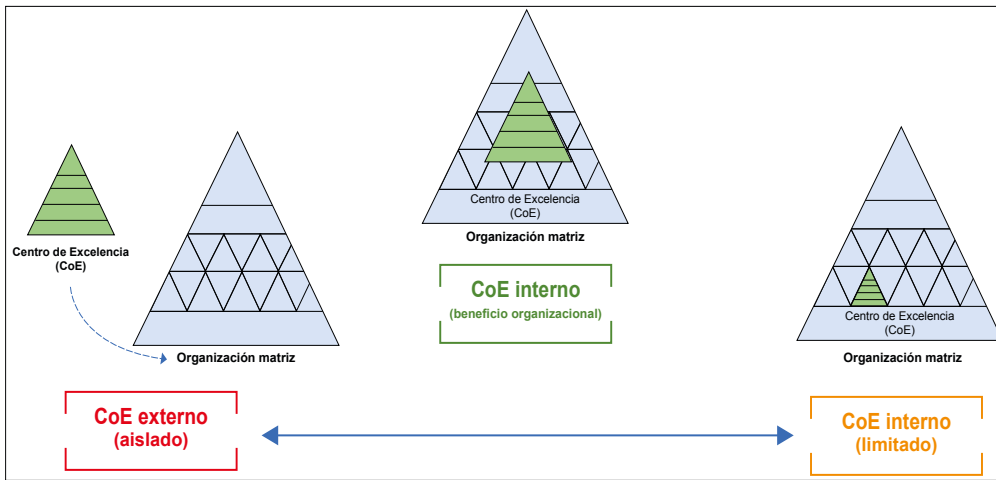


Figura 2. Ejemplos de CoE aislado, balanceado y limitado

Elaboración propia

La propuesta es empezar con un cambio en la estructura organizacional que sea ordenado, progresivo y que tenga un punto de inicio adecuado. Es necesario iniciar el cambio en la estructura tradicional de una organización por un CoE que empiece a aportar valor de forma rápida y progresiva y que, a su vez, traiga consigo otras aristas como el *agile mindset* y buenas prácticas de gobierno corporativo de TI. De esta forma es más fácil hacer tangibles los beneficios de una transformación, especialmente si las diferentes áreas de negocio de la organización se ven beneficiadas por las iniciativas digitales del mismo CoE que está impulsando la TD. En otras palabras, se trata de generar un cambio orgánico, que no busque ser aceptado desde el inicio por todo el personal, sino que busque generar valor para todos y que, como consecuencia de esta generación de valor, toda la organización desee subirse a la ola de la TD.

### 2.3 El *agile mindset* como cultura

Hasta este punto se ha discutido como una correcta estructura organizacional, que tenga como gatillador un CoE para expandir ese cambio estructural a toda la organización, ayuda con el cambio de la cultura organizacional hacia un *agile mindset*. Ahora es momento de entrar en detalle sobre cómo el segundo bloque de la pirámide ayuda a alcanzar el objetivo de la TD.

El *agile mindset* es un tema de moda desde hace varios años y que ha ido evolucionando en el tiempo. Esta forma de pensar y de actuar va de la mano con la estrategia a seguir para alcanzar la TD. Ya que el propósito de este estudio no es entrar en detalle de lo que es el *agile mindset* ni de sus beneficios, se asume que el lector tiene los conocimientos mínimos indispensables sobre el tema.



Es sabido que la TD tiene que ver con generar un cambio a nivel organizacional, sin embargo, este cambio funciona también como un proceso adaptativo a los actuales tiempos turbulentos que las organizaciones atraviesan. Lo que se debe considerar es que este proceso adaptativo debe convertirse en un *status quo*. Es decir, empezar a aceptar y abrazar el cambio, ya que este será perpetuo. La famosa frase, “Lo único constante es el cambio”, de Heráclito, lo resume bastante bien. Para este nuevo (y obligatorio) *status quo*, las organizaciones necesitan poseer el *mindset* adecuado para evitar generar resistencia a algo que es inevitable: el cambio; es mucho más provechoso aceptarlo y utilizarlo como palanca para obtener una ventaja competitiva. Es aquí donde entra como protagonista el *agile mindset*. Sin embargo, muchas organizaciones cometen el error de iniciar una TD con este bloque, y como ya se ha explicado anteriormente, sin una estructura organizacional adecuada un cambio de mentalidad como este no puede ser soportado. El *agile mindset* puede manifestar todos sus beneficios dentro de una organización solo con una estructura organizacional adecuada. Sin embargo, un cambio estructural no sucede de la noche a la mañana. Es necesario que sea progresivo para así evitar una disrupción organizacional negativa. Es por este motivo que se propone plantar la semilla de la mentalidad ágil en un CoE que tenga la estructura organizacional necesaria y que se encuentre dentro de la organización matriz para conseguir generar un balance. Es decir, que funcione y genere iniciativas digitales para toda la organización que traiga beneficios a nivel vertical y que impulse esta nueva forma de pensar y actuar. En otras palabras, que evangelice la agilidad a todos los miembros.

Existe, además, una razón adicional del porqué el *agile mindset* debe ser soportado por una correcta estructura organizacional y que además es el punto de conexión con la implementación del bloque de GEIT. Mike Cottmeyer, CEO de la empresa Leading Agile (2016), explica que cuando se pretende generar una transformación ágil (en adelante, TA), las organizaciones que fracasan suelen iniciar su camino intentando generar el cambio a través de la cultura o las prácticas ágiles. Este es un punto en común con aquellas organizaciones que buscan alcanzar la TD de la misma forma errada. Pretender comenzar un cambio altamente disruptivo (como la TD o TA) iniciando por un cambio cultural. Por otro lado, aquellas organizaciones que pretenden alcanzar el cambio hacia el *agile mindset* a través de la implementación de prácticas ágiles como Scrum o Kanban, cometen un error similar. Las prácticas ágiles deben ser soportadas por una estructura organizacional que les otorgue el nivel de autoridad, flexibilidad y sinergia suficiente con el negocio para poder, con el tiempo, evolucionar estas prácticas ágiles y generar equipos maduros que permitan expandir el *mindset* a toda la organización. Es virtualmente imposible aplicar estas prácticas ágiles sin tener una cultura ágil implementada. Las personas simplemente no verán el valor en las prácticas ágiles y empezarán a aplicarlas mal o no aplicarlas para nada.

Adicionalmente, Mike Cottmeyer recomienda que para alcanzar una correcta TA es necesario romper un poco el esquema tradicional de lo que es ágil. Este esquema tradicional indica que los roles ágiles son críticos para poder realizar la TA. Sin embargo, él sugiere que

se debe ir más allá de estos roles y comenzar a enfocarse en los beneficios que los roles traen. Estos beneficios se alcanzan a través de tres variables críticas en la gestión de proyectos ágiles. El *backlog*, los equipos y el *software* funcional y probado. Cuando se entra en el detalle de los beneficios de cada una de estas variables es cuando logramos encontrar la conexión entre una correcta estructura organizacional, un buen gobierno corporativo de TI y una cultura basada en *agile mindset*. Para esto analizaremos dos de las tres variables.

Empecemos por el *backlog*, que es básicamente un listado de los requerimientos de la iniciativa en curso. Si se hace de forma correcta, el fin principal del *backlog* es generar claridad para el equipo respecto a lo que se desea construir. Esta claridad a su vez genera un propósito para el equipo, que es una variable relacionada al aspecto motivacional y que impacta la productividad y enganche del equipo con la solución que están construyendo. Este aspecto de claridad conlleva a un aspecto de gobierno, que es tener un *oversight* de las múltiples iniciativas a través de la transparencia en la gestión de estas y que está soportado por el primer aspecto de claridad. El concepto de gobierno será explicado más adelante. Esta es la conexión que une el bloque *agile mindset* con el bloque GEIT. El GEIT debe tener claridad y visión de lo que sucede en los proyectos ágiles.

La segunda variable, el equipo, conlleva a la rendición de cuentas de este, ya que se trata de un equipo autoorganizado y que no solo es responsable de sus resultados, sino que responden por estos (*accountability* es el término en inglés). El aspecto de rendición de cuentas se traduce, finalmente, en la autonomía que posee un equipo de proyectos para poder ejercer sus funciones en un ambiente ágil. Es decir, la autonomía que posee el equipo los motiva e impulsa a hacer suya la solución que están creando. De esta forma se genera mucha más motivación y enganche para generar un resultado superior. Esta autonomía de equipo debe ser soportada por una estructura organizacional que les permita realizar estas funciones de forma idónea y sin bloqueos que provengan de otras partes de la organización. Esta es la conexión que une el bloque *agile mindset* con el bloque de la estructura organizacional. Los equipos de proyectos ágiles deben poder cumplir sus funciones en un ambiente adecuado, para lograr esto es necesario que las organizaciones reconozcan formalmente su autoridad a través de una estructura organizacional que apoye las prácticas, autoorganización, valores y principios. La figura 3 muestra la cascada de ideas y la conexión final entre bloques.

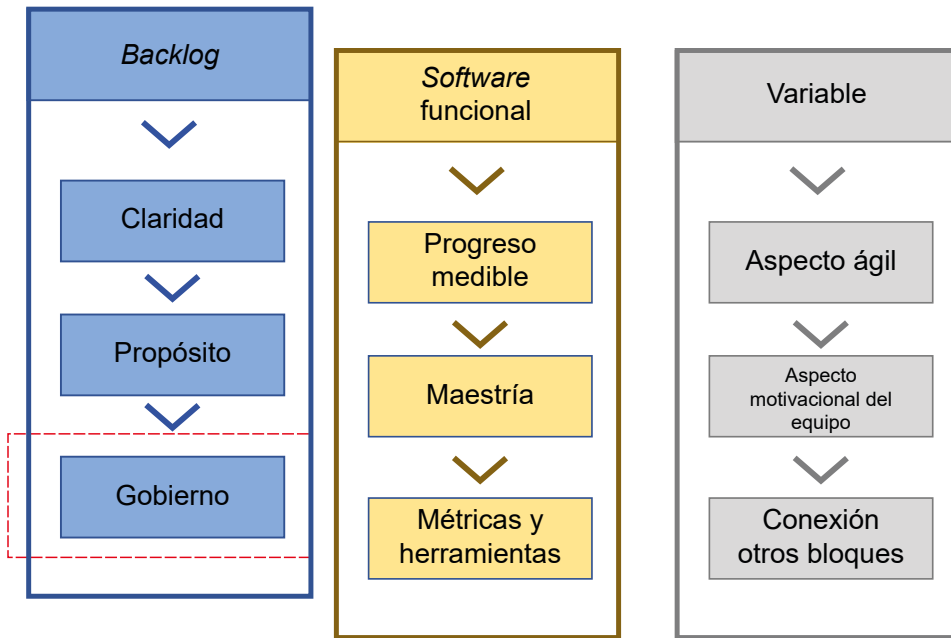


Figura 3. Variables ágiles y conexión entre bloques de la pirámide  
Elaboración propia

## 2.4 El gobierno corporativo de TI

El último y tercer bloque de la pirámide propuesta es el bloque del gobierno corporativo de TI (GEIT). El GEIT involucra un espectro de conocimientos bastante amplio y complejo. Uno que muchas organizaciones no entienden por completo y que, por lo tanto, no valoran como deberían. Una estructura organizacional correctamente implementada permite soportar todo lo que implica el generar un correcto GEIT; como políticas, procesos, roles, sinergias e incluso cambios adicionales en la estructura organizacional. No puede existir el bloque del GEIT sin el bloque base de la estructura organizacional ni tampoco sin el bloque del *agile mindset*.

En el contexto de la TD, el GEIT es una variable más que importante para poder alcanzar el objetivo de subirse a la tan codiciada ola transformacional.

Antes de discutir el porqué de su importancia, es necesario terminar de comprender el porqué este tercer bloque no puede existir sin los dos anteriores. La figura 1 muestra la pirámide y como estos tres bloques se unifican. Las formas de estos tres bloques tienen un propósito más allá de distinguirlos por colores. El bloque GEIT no podría mantenerse erguido y en su lugar sin el bloque de la estructura organizacional como base y sin tener el bloque *agile mindset* como soporte lateral. El implementar un GEIT en una organización es de por sí una tarea bastante

compleja. Al igual que sucede con la TA que discutimos anteriormente, el pretender iniciar la implementación de un GEIT desde un aspecto diferente al de la estructura organizacional tendría como resultado, casi de forma segura, un fracaso. Sobre todo, en el caso del GEIT y tal vez más crítico que en una TD, la estructura organizacional es un aspecto crítico ya que un GEIT bien implementado se traduce en cambios organizacionales no solo de *mindset*, sino de niveles de autoridad, flexibilidad, políticas, gestión de portafolios, programas y proyectos, arquitectura empresarial, ciberseguridad, procesos, entre otros aspectos (ISACA, 2015).

Por el lado del *agile mindset*, el GEIT puede apoyarse bien en algunos principios y valores de este, pero la sinergia se genera más por el contexto actual en el cual funcionan las organizaciones. El desconocimiento de qué pasará el día de mañana. La incertidumbre es una variable que, hoy en día, muchas organizaciones se ven obligadas a enfrentar, tanto desde el punto de vista externo (político, económico, social, etc.) como interno (falta de una visión clara de la solución que se desea construir, cambios organizacionales que pueden impactar una iniciativa, variables políticas internas entre áreas de la organización, etc.). Esta incertidumbre se traslada a la gestión de proyectos, que a su vez escala a la gestión de programas y portafolios de la organización, generando dificultades para gestionarlos de forma eficiente. Para contrarrestar las consecuencias negativas de la incertidumbre, el *agile mindset* le otorga a la organización las herramientas para poder disminuir su impacto y generar de la incertidumbre una ventaja. Es posible afirmar que la sinergia entre estos dos bloques se da más por un tema de contexto actual, que por una necesidad inherente para que una u otra sobreviva.

Es posible que una organización tenga implementado o implemente un GEIT sin siquiera saber el significado de lo que es el *agile mindset*. El valor que genera el tener un GEIT es indiscutible. Su complejidad va de la mano con la gran cantidad de beneficios que trae. Sin embargo, estos beneficios pueden ser mayores si se logra generar una sinergia con el *agile mindset* y con una estructura organizacional correcta.

Ahora que se explicó el contexto de las sinergias del bloque GEIT con el resto de la pirámide, es momento de explicar la razón de ser de este para la TD. Líneas arriba se mencionaba lo que implica implementar un GEIT y, si bien existen múltiples beneficios organizacionales que surgen de esto, solo se discutirán aquellos relacionados a impulsar o facilitar la TD en una organización. Estos beneficios son cuatro (ISACA, 2015).

- a) Alinear los objetivos estratégicos corporativos del lado del negocio con los objetivos estratégicos de TI.
- b) Disminuir el riesgo y costo de las iniciativas de TD.
- c) Evitar el desperdicio de oportunidades que generen valor.
- d) Coordinar de forma óptima las iniciativas digitales que surjan de un CoE.

El primer beneficio es el *core* de un GEIT. El alineamiento de las estrategias de negocio con las estrategias de TI es algo que puede sonar básico, común y hasta algo que podría ser considerado de sentido común. Sin embargo, los casos reales donde un GEIT funciona de forma óptima en organizaciones son muy pocos. Esto debido a la alta complejidad de implementación. Este alineamiento es indispensable si se desea alcanzar la cúspide de la pirámide propuesta en esta investigación. Pues, en la TD, si bien no se trata de tecnología sino de una estrategia, la tecnología está más que presente y sirve como una palanca que impulsa esta transformación. El área de TI tiene mucho que ver en esta y es indispensable que las personas involucradas tengan claridad de las estrategias corporativas de negocio.

El segundo beneficio, la disminución del riesgo y costo de las iniciativas de TD viene como resultado del ordenamiento que trae consigo el GEIT. No hablaremos de control, pues esto va un poco en contra del *agile mindset*. Más bien hablaremos de transparencia, la cual debe ser aplicada en la gestión de los proyectos, programas y portafolios de las iniciativas de TD que se lleven a cabo. Esta transparencia y *oversight* de las iniciativas traerá como consecuencia positiva la disminución de riesgos en estas y, por ende, la reducción del costo tras evitar estos riesgos.

El tercer beneficio, evitar desperdiciar oportunidades que generen valor, está relacionado con el concepto de P3 Governance (ISACA, 2015), o gobierno de las 3P (proyectos, programas y portafolios). Se refiere a la creación de un comité que permita filtrar de forma eficiente las iniciativas de TD que puedan surgir durante este proceso. Si un CoE fue implementado en una organización de forma balanceada, es decir, sin límites y cerca al negocio, y además empieza a tener éxito, es inevitable que casi todas las áreas de negocio de una organización empiecen a generar solicitudes de iniciativas digitales para que el CoE las ejecute. Esto conlleva a una lluvia de solicitudes de iniciativas que sería muy difícil de manejar, no solo por la cantidad y tiempo que implica analizar estas, sino por el incremento de la posibilidad de dejar de lado aquellas que realmente podrían traer valor a la compañía. Para evitar esta engorrosa situación, el GEIT propone la creación de pautas, procesos, métricas y roles que se encarguen de filtrar, ordenar y estructurar aquellas propuestas de proyectos digitales para que sean fácilmente digeridos por el CoE. Esto permitirá considerar solo aquellas iniciativas que estén correctamente estructuradas y tengan mayores probabilidades de generar valor para la organización.

Finalmente, el coordinar de forma óptima las iniciativas digitales va muy de la mano con los aspectos descritos en el párrafo anterior, y es que esto es también una consecuencia de tener transparencia, *oversight* y orden al momento de escoger qué iniciativas se llevarán a cabo. El GEIT abarca buenas prácticas para gestionar eficientemente los programas y portafolios organizacionales y con ellos asegurarse que las sinergias que puedan existir entre todos los proyectos generen el mayor valor posible.

## 2.5 Los cinco dominios de la transformación digital

Ahora que se han explicado a detalle los tres bloques que conforman la pirámide que permitiría alcanzar la TD, es momento de explicar los cinco dominios de la TD (Rogers, 2016) y como estos se alinean con los tres bloques de la pirámide descritos anteriormente.

La descripción de cada uno de los cinco dominios que se nombran a continuación no es exhaustiva. Se recomienda consultar el libro, *The Digital Transformation Playbook: Rethink Your Business for the Digital Age*, de David L. Rogers, para mayor detalle. Los cinco dominios de la TD, tal como los presenta Rogers, son: clientes, competencia, *data*, innovación y valor, representados en la figura 4.

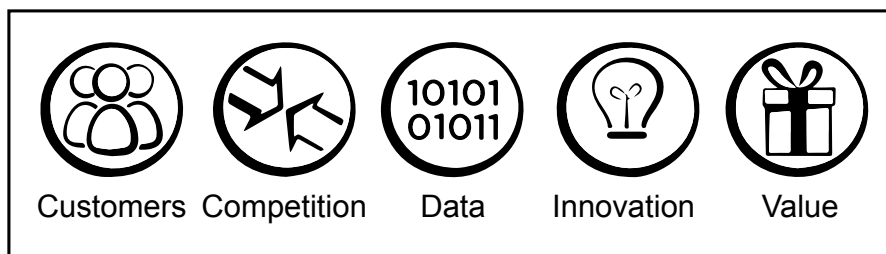


Figura 4. Los cinco dominios de la transformación digital (Rogers, 2016)

### 2.5.1 Clientes

El contexto actual requiere que las organizaciones piensen en sus clientes ya no como objetivos de venta sino como nodos en una red conectada que puede generar un alto valor para la organización, ya sea como un canal más de ventas que proporcione un retorno tangible o como un beneficio de imagen de marca que conlleve a ventas indirectas. Sin embargo, esta misma red implica que las organizaciones deben cuidar y mejorar la relación directa con sus clientes, ya que puede traer problemáticas a gran escala que se vuelvan exponenciales por el mismo poder de la red de clientes que les da beneficios. Es aquí de donde nace la necesidad de ser *user-centric* o tener un enfoque centrado en el usuario. No solo es imperativo tener el conocimiento para poder identificar que es lo que desean los clientes, sino también el cómo, cuándo y dónde.

### 2.5.2 Competencia

Hoy en día la competencia entre organizaciones ya no es binaria. De hecho, muchas organizaciones están generando colaboración con su competencia directa por retos mutuos que tienen dentro de su misma industria, o por el ingreso de nueva competencia con modelos de negocio totalmente disruptivos a los estándares de una industria y que los fuerza a tener que colaborar para sobrevivir. Sin embargo, las empresas también se han dado cuenta que muchas veces es

beneficioso el colaborar con un rival sin tener la presión de un competidor disruptivo. Uno de los ejemplos más conocidos de colaboración entre competidores es el de la división de teléfonos móviles de Samsung con su máximo rival, Apple. Samsung provee de las pantallas para casi todos los modelos de celulares de alta gama de Apple. Son los principales rivales de su industria, pero son también colaboradores entre ellos para mantenerse por encima de los demás. Ya no se trata solo de “competidor” o “no competidor”, se ha vuelto más complejo que eso. Es necesario buscar sinergias incluso con aquel rival que nadie pensaría podría ser un aliado.

### 2.5.3 Data

La *data* se genera hoy en día en cantidades inimaginables. Tanto *data* estructurada como no estructurada. Hasta hace veinte años, el reto era obtener y almacenar la *data*, hoy en día el reto es generar valor de la *data*, la cual existe en tal cantidad que es difícil poder utilizarla. Las organizaciones que posean las capacidades necesarias para poder explotar esta enorme cantidad de *data* tendrán una clara ventaja competitiva en su industria. Sin embargo, para esto es necesario también traer o desarrollar el talento que pueda generar las estrategias que permitan hacer una óptima minería de datos y obtener información y conocimiento que puedan traer valor a la organización.

### 2.5.4 Innovación

Este dominio describe el proceso por el cual las ideas se desarrollan, se prueban y se llevan al mercado a través de un proceso de prueba y error que se centra en un aprendizaje continuo a través de una rápida experimentación. Es decir, se utiliza lo que se conoce como producto mínimo viable (MVP, por sus siglas en inglés). La idea es fallar rápido y barato, no tenerle miedo al fracaso y ser capaz de rescatar lecciones aprendidas en el camino. En este punto, el *agile mindset* encaja bien. No solo por la forma de pensar y actuar que propone, sino también por las buenas prácticas de procesos propios de gestión de proyectos que ayudan a generar un MVP de forma óptima.

### 2.5.5 Valor

Este dominio se enfoca en mantener una propuesta de valor siempre evolucionando. Está relacionado al nuevo *status quo* que las organizaciones deben generar, aquel relaciona al cambio constante que les permite adaptarse a los nuevos retos y estar siempre un paso más allá de lo que el mercado exige. En vez de esperar y adaptarse cuando un cambio se vuelve de vida o muerte, las organizaciones necesitan enfocarse en tomar oportunidades emergentes de cambio y evolución. Para poder generar este valor, es necesario que se desarrollen habilidades que impulsen la evolución de la propuesta de valor de la organización hacia sus clientes.

El cómo los tres bloques de la pirámide se alinean con los cinco dominios de la TD (Rogers, 2016) descritos párrafos arriba serán descritos más adelante con un ejemplo real.

## 2.6 Alineamiento en la vida real

La TD es una estrategia para implementar a lo largo del tiempo hasta que esta se vuelva un nuevo *status quo* en la organización. En el contexto actual, los líderes deben tener la capacidad de reimaginar y reinventar el negocio para lograr adaptarlo a los cambios que vengan. Para lograr esto, es crítico poder alinear la agilidad organizacional y el GEIT con los dominios antes descritos. Pero esto de nada servirá sin una correcta estructura organizacional que soporte todo esto.

Para ilustrar mejor como se alinean los tres bloques de la pirámide con los cinco dominios de la TD, se utilizará como ejemplo la empresa de consumo masivo más grande del Perú y una de las más grandes de Latinoamérica, Alicorp. En enero del 2019, Alicorp lanzó oficialmente la nueva Dirección de Transformación Digital. Para ello, se decidió que era necesario pensar de una forma distinta y acorde a lo que el reto de la TD exigía. Es así como decide que esta nueva dirección funcione como un Centro de Excelencia Digital (CoE Digital), lo cual implicaba traer nuevos talentos para roles que jamás habían existido dentro de la organización. Estos roles fueron cubiertos, en su mayoría, por personal externo a la compañía que traía consigo nuevas capacidades, experiencia y conocimientos nuevos, así como por personal existente en la organización que sumarían su conocimiento actual del negocio, variable crítica para poder generar el impacto que se buscaba, y por ende el valor para la empresa.

A estos roles los llamaron *leads*, que serían las cabezas de una estructura organizacional nueva, empoderada y que permitiría construir sobre esta las capacidades de los otros dos bloques de la pirámide, *agile mindset* y GEIT. Los roles que los *leads* asumieron fueron: *experience design lead*, *advanced analytics lead*, *automation lead*, *technology lead*, *agile lead* y *digital transformation office lead*. Estos seis roles generarían sinergias entre ellos para complementarse y estarían alineados con los cinco dominios de la TD como se muestra en la figura 5.

Alicorp tuvo la visión de crear un CoE Digital con un nivel de autonomía sin precedentes para permitirle generar su propia estructura organizacional incorporando roles que antes no existían dentro de la empresa. Esta estructura se encuentra dentro y forma parte de la organización matriz (Alicorp), y existe para servir a todos los niveles organizacionales. De esta forma el CoE Digital no se encuentra aislado del negocio, por lo que se pueden generar mejores sinergias con las diferentes vicepresidencias, direcciones y equipos de la organización. Por esta misma razón, las iniciativas digitales que nacen del CoE no están limitadas, pues no se enfocan en una sola área de negocio, sino más bien a todas las áreas existentes de la compañía, permitiendo de esta forma asegurar una participación de los diferentes *stakeholders* del negocio en los proyectos digitales que ellos mismos solicitaron y en aquellos que nacieron por iniciativa del CoE Digital.



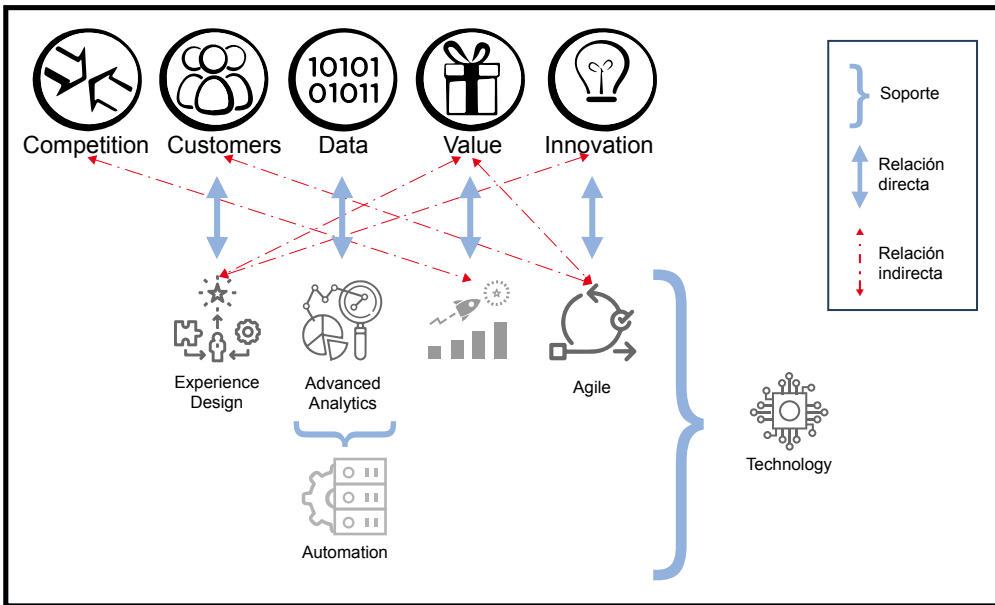


Figura 5. Sinergias y alineamiento del CoE Digital de Alicorp con los cinco dominios de la TD  
Elaboración propia

Esta estructura facilitó a su vez la generación de un *agile mindset* (el segundo bloque de la pirámide), el cual se convirtió en *status quo* dentro del CoE Digital y que permitió sembrar la semilla de agilidad organizacional que ya se encuentra expandiéndose al resto de Alicorp. En paralelo a la evolución propia del *agile mindset* dentro de la organización, la generación de un GEIT se está llevando a cabo para la implementación de las políticas, procesos, roles y demás aspectos relacionados al gobierno de TI.

Un aspecto de la TD, que podría ser catalogado como sentido común, pero que en la realidad usualmente es descartado, es que todo el proceso de lograr la TD es progresivo, toma tiempo, y es necesario iniciarlo por etapas. Esta es una de las razones por las que se decide que el viaje de esta transformación sea iniciado y llevado a cabo por un equipo específico con un empoderamiento que vaya de la mano con lo que exige este tipo de iniciativa. La iniciativa de la TD nace, se desarrolla, evoluciona y se expande al resto de la organización desde el CoE Digital. La pirámide propuesta debe ser aplicada en una escala lo suficientemente pequeña para ser gestionada de forma eficiente, pero con la suficiente autonomía, autoridad y capacidades para empezar a generar el valor para la organización. De esta forma se generará el impulso necesario para la adopción orgánica de los pasos propuestos en este estudio. Es así como el camino a la TD se vuelve una evolución natural para la organización.

La pirámide debe ser aplicada en una escala menor (CoE Digital), para luego llevarla a una escala organizacional (Alicorp). Una vez que las buenas prácticas, procesos, roles, estructura y conocimientos estén asentados en el CoE Digital, y que exista un alineamiento de los tres bloques con los cinco dominios de la TD, es en ese momento que el nuevo *status quo* del CoE Digital estará completo, y el camino hacia la TD asegurado.

Esto no quiere decir que mientras el CoE Digital esté en este proceso evolutivo no pueda empezar a pavimentarse el camino del resto de la organización hacia el objetivo transformacional. De hecho, es todo lo contrario. El CoE Digital ha impulsado desde su concepción y seguirá impulsando esta transformación a todo nivel organizacional. No es necesario tener una pirámide completa para recién pretender iniciar la transformación en el resto de la organización. El negocio no puede parar, y jamás lo hará, ni siquiera por la TD.

## 2.7 Los agentes transformacionales

Hay una última variable que no ha sido mencionada, pero que es crítica para poder impulsar todo lo propuesto en este artículo. La gente.

Los roles de aquellas personas que se encargan de generar, impulsar y ejecutar esta transformación ya fueron descritos: los *leads* del CoE Digital. Pero el título complementario, asignado a estos líderes para fines del presente estudio, es el de agentes transformacionales.

Ellos son quienes lideran todos los frentes del equipo encargado de generar la TD en la organización. Sus conocimientos, experiencia y sinergias serán las que definan el éxito o fracaso de una iniciativa tan grande como esta. Es así como se vuelve sumamente crítico seleccionar a aquellos que lideren con pasión, vocación de servicio y posean un enfoque humano que abarque no solo buscar beneficios para su equipo inmediato o accionistas, sino que incluya a los clientes, a los proveedores y a las sociedades donde su organización genera al menos un mínimo impacto.

El lugar en la pirámide donde estos agentes actúan es estratégico, ya que deberán asegurar las sinergias de los tres bloques de la pirámide y aquellos dominios asociados al desarrollo de toda la pirámide. Es así como su posición es crítica para el éxito. La figura 6 muestra su posición.

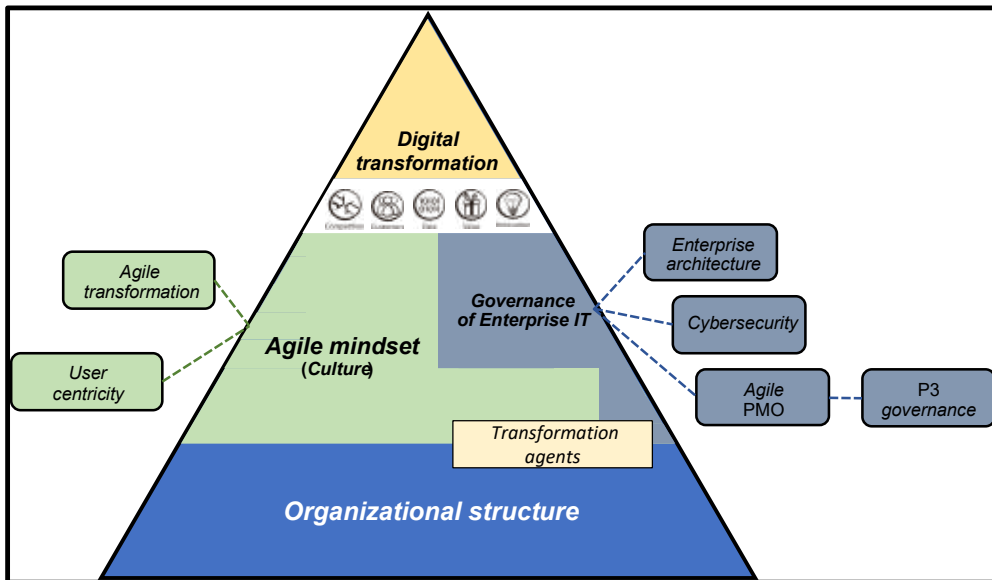


Figura 6. Los agentes transformacionales en la pirámide

Elaboración propia

Es este equipo de líderes quienes guían a toda una organización a encontrar nuevas formas de hacer las cosas y que, en el camino, disfrutan haciéndolo.

### 3. RESULTADOS

El camino y la pirámide propuesta para alcanzar el objetivo de la TD en una organización están basados, en gran parte, en el estudio de casos de éxito, teorías, marcos de trabajo y propuestas teóricas de diversas fuentes. Sumado a esto, la experiencia vivencial de participar de una estrategia viva de la más grande empresa de consumo masivo del Perú, la cual supo crear y alinear los tres bloques, de estructura organizacional, *agile mindset* y GEIT, ejemplifica bien todo lo propuesto.

A junio de 2019, con apenas seis meses de existencia, el CoE Digital de Alicorp había generado tal impacto en la organización, que su meta de incremento recurrente del EBITDA (beneficios antes de intereses, impuestos, depreciación y amortización, o EBITDA, por sus siglas en inglés) organizacional se incrementó en un 512 %. Lo cual no fue resultado de un mal cálculo inicial o mala estimación del impacto en el EBITDA, sino más bien del excelente desempeño que el CoE Digital demostró con las capacidades recién adquiridas, apalancadas por una estructura organizacional única y un *agile mindset* alineado a los objetivos estratégicos de la organización.

Actualmente, las sinergias de los seis *leads* del CoE Digital se plasman en la figura 7 que ilustra las sinergias y el proceso por el cual pasan casi todas las iniciativas planeadas hasta el 2021.

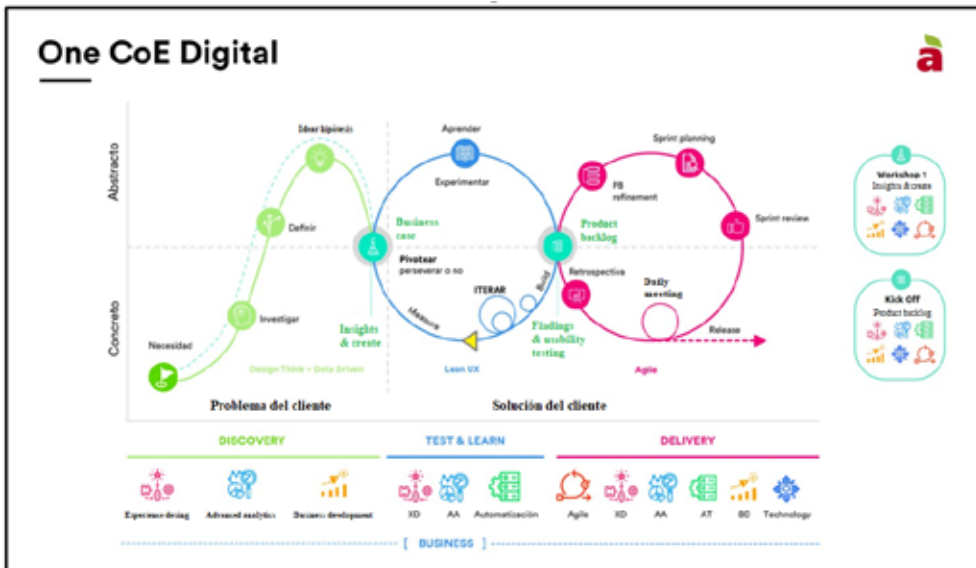


Figura 7. Sinergias del CoE Digital de Alicorp  
Fuente: CoE Digital de Alicorp

Los resultados de iniciar una travesía de TD no serán inmediatos. Se trata de un proceso largo, muchas veces doloroso y con un camino plagado de retos.

#### 4. CONCLUSIONES

Para subirse a la ola de la TD es necesario:

- Todo viaje de TD debe iniciar con los altos ejecutivos de la organización dispuestos a impulsar, participar y comprometerse con el cambio.
- Iniciar el proceso de TD con la estructura organizacional adecuada, la cual puede implementarse dentro de un CoE que sirva como punto de partida para el resto del cambio organizacional. Sin embargo, se debe considerar que este CoE debe tener el nivel adecuado de autonomía, autoridad y flexibilidad para llevar a cabo su labor.
- Generar un *agile mindset* apalancado en una estructura organizacional adecuada que le brinde la autonomía, autoridad, flexibilidad, cercanía con el negocio y sinergias necesarias para generar un cambio (progresivo) orgánico en el *mindset* de todas las personas de la organización.
- Implementar un GEIT a nivel del CoE que luego pueda ser elevado a toda la organización a través de los cuatro beneficios relacionados a la TD. Si el resto de la

organización empieza a ver los beneficios de un GEIT correctamente implementado, esto servirá como palanca para llevarlo a un mayor nivel organizacional.

- Asegurarse de que el CoE que se genere sea balanceado y evitar caer en los extremos, tales como el CoE aislado o el CoE limitado. De esta forma, el crecimiento y aceptación del CoE será orgánico, ya que las áreas de negocio empezarán a ver el valor que se genera de una estructura adecuada y correctamente posicionada dentro de la organización.
- Asegurarse de alinear correctamente las capacidades del CoE con los cinco dominios de la TD propuestos (Rogers, 2016). Las capacidades deben permitir alcanzar y desarrollar cada uno de estos dominios.
- Es posible iniciar una transformación ágil, pero esta debe estar alineada con el objetivo final de alcanzar la TD. De lo contrario los recursos podrían destinarse erróneamente.
- Formar un equipo de agentes transformacionales que estén a la altura del reto. Es necesario que cuenten con experiencia en sus áreas de conocimiento, pero es mucho más crítico que se asegure las sinergias entre todos los líderes de los frentes de la TD.

## REFERENCIAS

- Harshak, A., Schmaus, B., y Dimitrova, D. (2013). *Building a digital culture: How to meet the challenge of multichannel digitization*. USA: Booz & Company, PricewaterhouseCoopers.
- Herbert, L. (2017). *Digital Transformation: Build Your Organization's Future for the Innovation Age*. London: Bloomsbury Publishing.
- ISACA. (2015). *CGEIT Review Manual, 7th Edition*. Schaumburg, IL, USA: ISACA Knowledge Center, ISACA.
- Leading Agile. (2016). Why Agile Fails in Large Enterprises - Large Scale Agile Transformation [archivo de video]. Recuperado de <https://www.youtube.com/watch?v=Oo3zlOTbN2E>
- Project Management Institute. (2017). *Agile Practice Guide*. Newtown Square, Pennsylvania, USA: Project Management Institute, Inc.
- Project Management Institute. (2016). *Governance of Portfolios, Programs and Projects: A practice guide*. Newtown Square, Pennsylvania, USA: Project Management Institute, Inc.
- Rogers, D. L. (2016). *The Digital Transformation Playbook: Rethink your business for the digital age*. New York: Columbia University Press, Columbia Business School.

Siebel T. M. (2019). *Digital Transformation: Survive and Thrive in an Era of Mass Extinction*. New York: RosettaBooks.

Westerman, G., Bonnet, D., y McAfee, A. (2014). *Leading Digital: Turning Technology into Business Transformation*. Boston: Harvard Business Review, Harvard University.

# Sistema de detección y clasificación de peces utilizando visión computacional

Rodrigo Mejía  
20140809@aloe.ulima.edu.pe / Universidad de Lima, Perú

Gianfranco Rosales  
20141199@aloe.ulima.edu.pe / Universidad de Lima, Perú

Recepción: 21-6-2019 / Aceptación: 8-8-2019

**RESUMEN.** La gestión de los recursos hidrobiológicos implica tanto el aspecto ecológico a través del equilibrio del ecosistema, como el aspecto económico mediante el control de la cantidad y calidad de los recursos pesqueros producidos en el Perú. En la actualidad, labores relacionadas a esta gestión son realizadas por empresas privadas y entidades del Estado como el Imarpe. La misión de estas es proteger la calidad de los recursos que llegan a los hogares de millones de peruanos. Esta investigación busca desarrollar un sistema para la detección, clasificación y, finalmente, la medición de diversas especies de peces, utilizando técnicas de visión computacional como el algoritmo SURF y redes neuronales convolucionales. Las pruebas, utilizando dos especies de peces, demostraron que la identificación alcanza un nivel de precisión del 90 % y que la clasificación alcanza una precisión del 80 %. Estos valores se obtienen bajo determinadas condiciones que se comentan en el desarrollo del artículo.

**PALABRAS CLAVE:** visión computacional, SURF, *background subtraction*, *border detection*, red neuronal convolucional, *machine learning*, detección, clasificación, inceptionV3, peces, CNN

## Fish Detection and Classification System Using Computational Vision

**ABSTRACT.** The management of hydrobiological resources involves both the ecological aspect through the balance of the ecosystem, and the economic aspect through the control of the quantity and quality of the fishery resources produced in our country. Currently, work related to this management is carried out by private companies and state entities such as Imarpe. Their mission is to protect the quality of the resources that reach the homes of millions of Peruvians. This research aims to develop a system for the detection, classification and finally measurement of various species of fish, using computational vision techniques such as the SURF algorithm and convolutional neural networks. The tests, which used two fish species, showed that the identification reaches a 90% accuracy level and the classification reaches an 80% accuracy level. These values are achieved under certain conditions that are discussed in the article.

**KEYWORDS:** computational vision, SURF, background subtraction, border detection, convolutional neural network, machine learning, detection, classification, inception V3, fish, CNN

## 1. INTRODUCCIÓN

Los recursos marinos representan una fuente de recursos valiosa para el desarrollo de la economía y la alimentación de sus comunidades tanto para el mundo como para nuestro país. La Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO, por sus siglas en inglés), indica que el consumo aparente de kilogramos per cápita para la alimentación y otros fines es cercano a la utilización total de recursos disponibles. Particularmente, nuestro país es uno de los países más ricos en recursos pesqueros del planeta, ocupando el quinto lugar de producción pesquera en el año 2016 (FAO, 2018). Durante el período enero-junio del 2018, el Perú exportó 839,3 toneladas de pescado, lo cual en términos monetarios significó 1787,6 millones de dólares (Sociedad Nacional de Pesquería, 2018).

Acorde al Codex Alimentarius de la FAO (CRFM, 2016), la trazabilidad es definida como “la capacidad de seguir el movimiento de un alimento a través de etapas específicas de producción, elaboración y distribución”. En el caso de instituciones como el IMARPE (Instituto del Mar del Perú), en donde realizan la clasificación y medida de especies de forma manual, le podría generar un ahorro de esfuerzo y tiempo automatizando estos procedimientos. En el caso de terminales pesqueros también podrían acceder a este ahorro, facilitando la trazabilidad debido al menor tiempo para analizar los peces.

Dichas estadísticas y antecedentes expresan la importancia de contar con recursos de calidad que, al ser extraídos del mar, presenten condiciones óptimas para la llegada al consumidor final. Por ello, en esta investigación se plantea un sistema de detección, clasificación y medida de determinadas especies de peces, tarea realizada manualmente en los laboratorios de instituciones como el Instituto del Mar del Perú y que, de desarrollarse adecuadamente, significaría un ahorro de esfuerzos importantes para este tipo de organizaciones.

## 2. ESTADO DEL ARTE

En esta sección se analizan trabajos relacionados a las aristas claves de esta investigación: la detección de objetos y la clasificación de imágenes mediante aprendizaje supervisado. Por un lado, la discusión sobre detección de objetos abarca trabajos relacionados en otros campos de estudio (p. ej., detección de ballenas) y la comparación de resultados del algoritmo SURF con otros afines como SIFT, ORB, entre otros. Acto seguido, para el análisis de los trabajos relacionados a la tarea de clasificación de imágenes, se discute la aplicación de redes neuronales convolucionales (también en su abreviatura como CNN) en la tarea de clasificación, así como las arquitecturas de las mismas, y aplicaciones e investigaciones orientadas al reconocimiento de peces.



## 2.1 Detección de objetos en otros campos de estudio

Las investigaciones relacionadas a la detección de objetos específicos vinculados a la visión computacional son las siguientes:

Nadarajan, Chen-Burger y Fisher (2011) diseñaron un sistema capaz de analizar vídeos bajo el agua para detectar, rastrear y contar peces. Para la detección de peces, los autores utilizaron el modelo gaussiano y el H-model que ahorra el costo computacional, donde los valores de precisión terminaron entre 80,5 %–89,5 %.

Mahajan y Perkins (2015) se enfocaron en detectar e identificar ballenas en un conjunto de imágenes tomadas de manera aérea. Para la detección de ballenas se utilizaron las técnicas *cascade object detector* (COD) y el método de agrupamiento de K-medias (KMC). Para la identificación de las ballenas utilizaron SIFT y *support vector machine* (SVM). Los resultados de esta investigación concluyeron que KMC detecta mejor que COD en el 90 % de los casos, mientras que no consiguieron que los métodos de identificación funcionen de manera satisfactoria.

Chincha y Tian (2011) propusieron un método de reconocimiento de objetos para ayudar a invidentes para que encuentren sus pertenencias mediante el algoritmo de visión computacional SURF. Los objetos por detectar eran llaves, teléfonos celulares, billeteras, lentes de sol y zapatillas. Este método propuesto por los autores terminó con un 84 % de precisión para los cinco tipos de objetos. Por otro lado, Shujuan, Zhize, Xingang, Guan, Wenqi y De (2015) utilizaron el mismo algoritmo para realizar una detección de vehículos a tiempo real. Sus resultados fueron positivos, solo con pequeños percances como la confusión entre el carril y el vehículo, y en condiciones de alta y baja luminosidad, donde el detector de vehículos no funcionaba de la manera esperada.

## 2.2 Reconocimiento de peces

Chen, Sun y Shang (2017) proponen un sistema de clasificación de peces automático para el etiquetado de los peces usando tecnologías de aprendizaje profundo, específicamente modelos combinados de tecnologías de redes neuronales convolucionales. La clasificación se trabajó con distintas arquitecturas combinadas como ResNet50, InceptionV3 y VGG16. La técnica que obtuvo un mejor resultado respecto a la función de error fue la técnica InceptionV3, entrenada con *data* externa con el valor de 0,853 de pérdida.

Frost y Geisler (2016) comparan los métodos de clasificación bayesianos ingenuos, máquinas de vectores de soporte y árboles aleatorios para clasificar especies de peces del mismo conjunto de datos trabajado por Chen, Sun y Shang (2017). La precisión de clasificación obtenida es de 7 %, 40 % y 80 % para cada método de clasificación, respectivamente. Posterior a la optimización del método de árboles aleatorios se logró obtener un valor de precisión general del modelo en el entrenamiento de 99,5 % y en la fase de prueba de 83,4 %.

Dada la revisión de literatura mencionada, el equipo de trabajo concluyó que las técnicas de redes neuronales convolucionales profundas y el algoritmo SURF son válidas para la propuesta de solución.

### 3. METODOLOGÍA

Esta sección detallará determinados tópicos para el entendimiento del nacimiento de la solución que se plantea. Primero, hace referencia a las etapas de la investigación. A continuación, hace referencia y detalla la estructura de la solución propuesta. Luego, se mencionan los métodos que se probaron durante el proceso. Finalmente, se explica la manera en que se validó el trabajo realizado.

#### 3.1 Etapas de desarrollo de la solución

Durante la primera fase, se buscaron métodos para resolver el problema planteado: poder detectar y clasificar peces en imágenes que presenten determinada escena. Para lograr esto, se investigó arduamente de manera que se identificaron métodos potenciales para el planteamiento de una futura solución: emplear algoritmos *hand-crafted* como lo son SURF, SIFT, etc., y el uso de redes neuronales convolucionales profundas para la tarea de clasificación de imágenes.

La siguiente fase o etapa consistió en la obtención de los datos (imágenes) a trabajar. Para el desarrollo de la prueba de concepto que se realizó, se empleó un conjunto de datos obtenidos de la organización Kaggle. Gracias a los optimistas resultados obtenidos, se propuso mejorar y ajustar la solución preliminar propuesta y trabajar con datos reales, imágenes que presenten a los peces que se encuentran en nuestro litoral.

Esto da pase a la tercera etapa: la obtención de los datos finales. Para esta labor, el equipo de trabajo se contactó con múltiples organizaciones. En última instancia, se recibió el apoyo del Instituto del Mar del Perú. Gracias al apoyo del ingeniero Julio Alarcón, jefe de Artes de Pesca de la institución, el equipo pudo programar una visita guiada a cargo del personal de Gestión de la Calidad del Terminal Pesquero de Ventanilla, a fin de capturar las imágenes que se emplean en el presente trabajo. Finalmente, como última etapa se encuentra la optimización de los métodos. Esta se realizó en base a los resultados empíricos obtenidos durante la fase de experimentación.

#### 3.2 Estructura de la solución

La solución propuesta posee determinados procesos que el sistema realizó para conseguir el objetivo planteado, donde se empezó con el preprocesamiento de las imágenes y se culminó

con la visualización e interpretación de los resultados de clasificación en la interfaz gráfica de usuario.

Se realiza el preprocesamiento de imágenes. Este consiste en dos tareas esenciales: el redimensionamiento de las imágenes y el entrenamiento realizado de manera iterativa, evaluando y analizando los resultados obtenidos de los valores de pérdida y precisión que se obtienen durante el proceso de entrenamiento y validación. La salida de este último mencionado es un archivo en formato HDF5 conteniendo los pesos y conexiones ideales generadas en el proceso de entrenamiento de la red. Este proceso fue realizado íntegramente en la nube de Google Colaboratory, por lo que se logró cruzar la barrera de recursos computacionales que se tuvo al comienzo de la investigación.

Paralelamente, se realizó la detección de características de una imagen base, donde el algoritmo SURF trabajó con una silueta característica de un pez. Posteriormente, se extrajeron las características de las imágenes reales, donde se quiso identificar y detectar al pez en un contexto real. Este último paso se logró gracias al uso de la técnica *background subtraction* y *border detection*, donde el primero se encarga de abstraer al pez del contexto total de la imagen y el segundo de detectar los bordes de la imagen. Tras la detección, se hizo el *feature matching* entre ambas imágenes, donde se analiza si la imagen del conjunto de datos tiene las características de la imagen base. Si la imagen las contiene, se procesa el pez detectado generando una imagen, la cual sirve de entrada a la red neuronal convolucional para poder realizar la clasificación en las imágenes. Finalmente, teniendo como entrada los pesos generados en el entrenamiento y la imagen del pez detectado, se clasifica. En la siguiente figura se muestra de manera gráfica el proceso descrito.

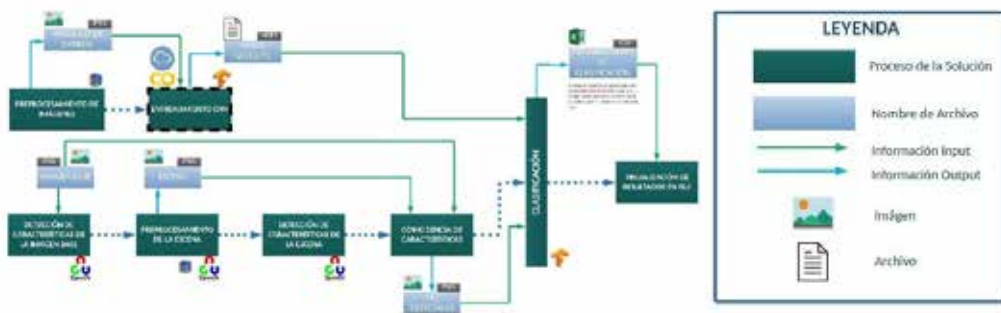


Figura 1. Estructura de la solución propuesta

Elaboración propia

### 3.3 Métodos probados

Respecto a la tarea de clasificación, se probaron tres arquitecturas de red neuronal convolucional conocidas. Estas fueron VGG-16, ResNet e InceptionV3. Se optó, finalmente, trabajar con la última mencionada por presentar los mejores resultados preliminares en base a las pruebas realizadas. Por ello, en esta investigación este modelo se optimiza a nivel de parámetros y agregación de capas con el fin de obtener las mejores conexiones y pesos para el modelo final. En la siguiente figura se observa gráficamente el modelo InceptionV3 con la agregación de capas en la parte superior. En la parte inferior se observan las últimas capas del modelo original de InceptionV3, las cuales son precargadas con los pesos de ImageNet. Por otro lado, las capas agregadas se entrenan con las imágenes del conjunto de imágenes de peces obtenidos por el equipo de trabajo.

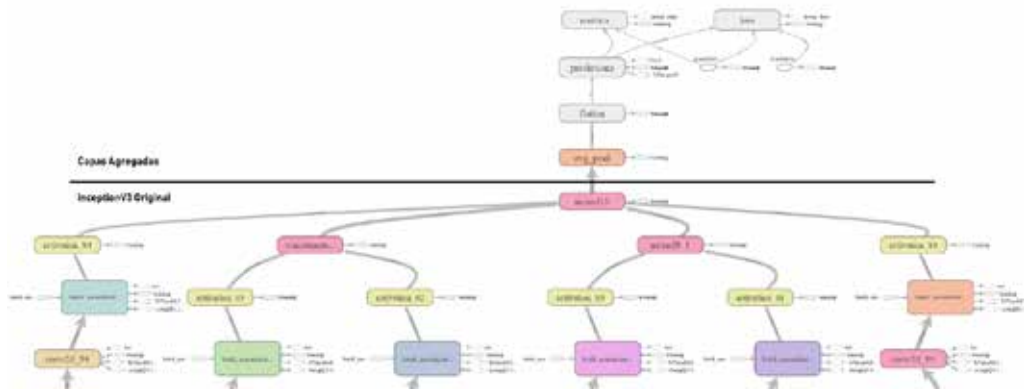


Figura 2. Modelo InceptionV3 adaptado a la propuesta  
Elaboración propia

Por otro lado, para la tarea de detección se probaron tres algoritmos de visión computacional: SIFT, SURF y ORB. Se optó trabajar con SURF porque detectó satisfactoriamente a los peces en sus respectivas escenas. Otras técnicas implementadas para poder facilitar la detección fueron *background subtraction* y *border detection*.

## 4. EXPERIMENTACIÓN

Se decidió utilizar un arco para poder realizar la toma de las fotos con los siguientes equipos computacionales: una cámara Logitech c922 Pro Stream que tiene como resolución máxima 1080p/30fps y un ángulo de visión de 78°, un ordenador de procesamiento remoto de Google Colaboratory (GPU Tesla K80 con 2496 CUDA cores) y un ordenador con las librerías instaladas.

#### 4.1 Detección del pez en la escena




Para la detección de los peces se utilizaron las siguientes técnicas: el algoritmo de visión computacional SURF, la técnica de *background subtraction* y *border detection*. Para la detección y caracterización de la imagen base, fue utilizado un contorno de un pez, el cual presenta las siguientes condiciones:

- Fondo blanco.
- Es únicamente un contorno genérico de un pez.
- La imagen tiene características que todos los peces del conjunto de datos poseen.

Para empezar, se procesó la imagen base, encontrando sus características mediante el algoritmo SURF. Posteriormente, se aplicó *background subtraction* en la escena del pez que se quiere detectar su presencia en la escena, donde se comparan dos imágenes: la de la escena sin el pez y la del pez en la escena. En las imágenes presentadas, se muestra el pez cojinova.

Tabla 1

*Imágenes a procesar mediante background subtraction (fondo y con la escena) y su posterior resultado*

Imagen			
Descripción	Imagen "fondo" sin el pez en la escena	Imagen con el pez en la escena	Resultado tras el <i>background subtraction</i>

Elaboración propia

Como se puede observar en la tabla 1, las imágenes tras el *background subtraction*, el pez se ve de manera clara. Posteriormente, se aplica el *border detection* para poder detectar el contorno de estas imágenes.

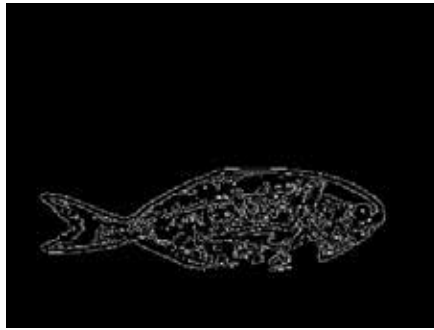


Figura 3. Resultado tras aplicar *border detection* en el resultado anterior

Elaboración propia

Debido a que se aplicó el *background subtraction*, previamente, el contorno del resto de peces en la imagen no se visualiza de manera tan clara como el del pez objetivo. Ya teniendo el borde detectado de ambas imágenes (la imagen a detectar y la silueta genérica) se procede a la etapa de coincidencia de características detectadas entre ambas imágenes.

Esta etapa de coincidencia de características (*feature matching*, en inglés) se realiza mediante FLANN (Fast Library Approximate Nearest Neighbor), la cual es una librería que permite comparar los puntos de características encontrados en ambas imágenes. Tras realizar la coincidencia de características, se pasa a la prueba de ratios de distancia entre los puntos vecinos para filtrar la similitud de coincidencias, donde se definen las coincidencias que van a ser tomadas en cuenta.

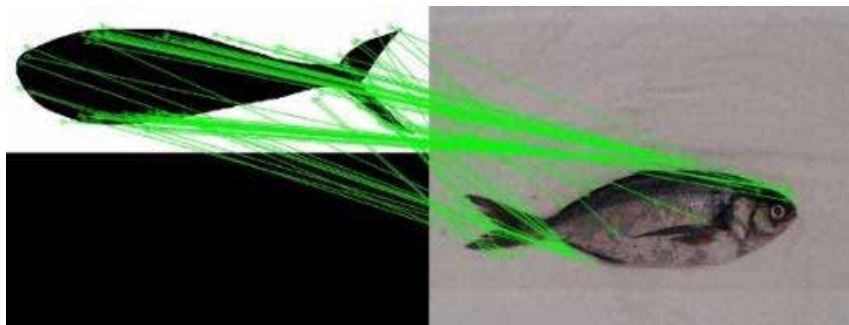


Figura 4. Resultados de la detección

Elaboración propia

Finalmente, la parte de detección del pez en la escena concluye devolviendo como resultado el pez detectado y extraído de la escena para su posterior clasificación.



Figura 5. Exhibición del *output* final de la etapa de detección

Elaboración propia

## 4.2 Clasificación de peces

Para la ejecución de los experimentos y pruebas relacionadas a la tarea de clasificación, se realizaron determinados pasos para lograr el entrenamiento exitoso en la nube de Google Colaboratory.

Siendo el objetivo subir las imágenes a la carpeta de trabajo en la nube de Colaboratory, primero se hicieron determinadas tareas de forma local. Las tareas realizadas se detallan en el siguiente orden:

- 1) Etiquetado de las imágenes para poder saber qué imagen corresponde a determinada especie. Esta tarea se realizó de forma manual.
- 2) Para las imágenes que contaban con el formato HEIC, se empleó un *software* de conversión de formato a JPEG.
- 3) División del conjunto de datos en una proporción de 80 % para entrenamiento y 20 % para validación.
- 4) Dadas las dimensiones originales de las imágenes: 4032 x 3024 píxeles, se redimensionaron las mismas con el objetivo de optimizar el tiempo del proceso de entrenamiento en la nube. Las nuevas dimensiones obtenidas fueron de 299 x 299 píxeles.

### 4.3 Entrenamiento de la red neuronal

Se realizaron numerosas iteraciones de entrenamiento caracterizadas por la variación de los parámetros de tamaño de lote, cantidad de épocas y el parámetro de guardado del modelo y sus pesos. El mejor resultado obtenido se dio en la iteración que presentó mejor convergencia en el tiempo de los valores de precisión y pérdida de entrenamiento y validación durante el proceso de entrenamiento. Los valores de los parámetros para la iteración mencionada son: tamaño de lote de 32, número de épocas 30 y ratio de aprendizaje de 0,001.

Los valores de precisión obtenidos durante esta fase son de 0,977 para el entrenamiento y 0,962 para la validación. Asimismo, los valores de pérdida y entrenamiento obtenidos son de 0,18, aproximadamente. Estos se pueden apreciar en la figura 6.

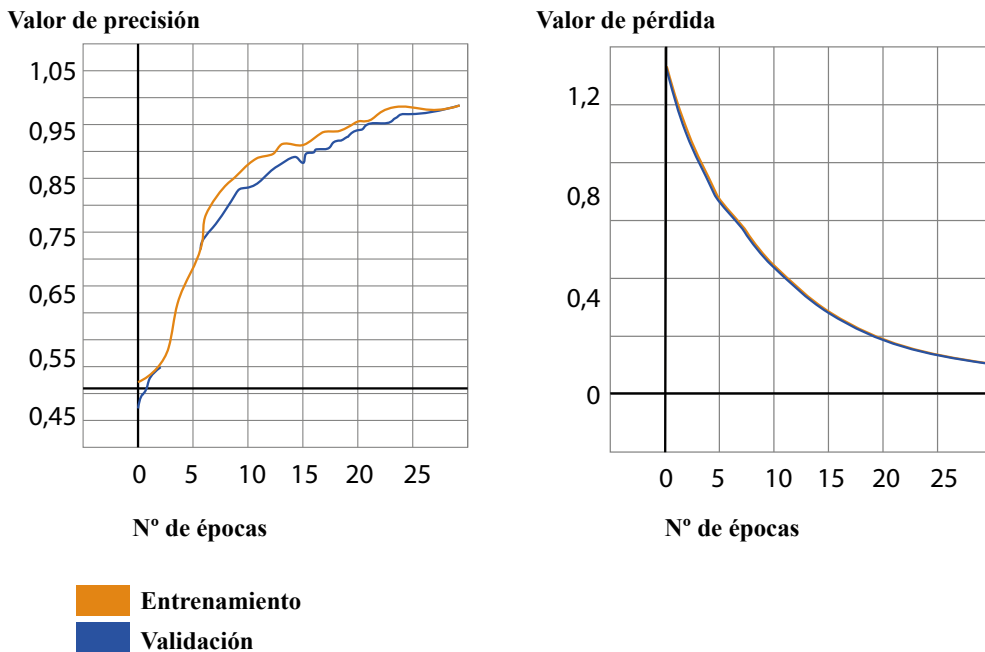


Figura 6. Valores de precisión y pérdida de entrenamiento y validación

Elaboración propia

## 5. RESULTADOS

Los hallazgos de este trabajo de investigación se presentan en dos etapas de experimentación: detección y clasificación; y en dos fases, la fase de pruebas con imágenes impresas de los peces y las pruebas realizadas con peces reales.



En la primera fase se obtuvieron los siguientes resultados:

Respecto a la detección, se procedió a detectar un conjunto de imágenes que tenía las siguientes características: un total de veinte imágenes, siete imágenes por cada especie y seis imágenes sin un pez en la escena. Se obtuvo lo siguiente:

- Tras la implementación de *background subtraction*, *border detection* y la conversión de la imagen a binaria (*thresholding*), el algoritmo detectó satisfactoriamente a los peces en la escena.
- Se detectó correctamente el 90 % de las imágenes.
- El recorte de la imagen, tras agregar ciertos valores de tolerancia (p. ej., si se detectó la cola de un pez de manera parcial, se recorta con un margen de tolerancia para evitar cortar la cola), funciona de la manera esperada.

Respecto a la clasificación, se definieron dos clases trabajadas, bonito y cojinova. La clase “no pez” no se incluyó, pues antes de la clasificación el sistema realizó la detección. Los ajustes en el conjunto de imágenes final con el que el modelo entrenó. Se trabajó con una cantidad homogénea y balanceada para cada clase. Esto significa que se entrenó al modelo durante 100 épocas, con ciento seis (106) imágenes por cada clase, se realizó la validación paralela con veintiocho (28) imágenes y se hicieron las pruebas finales con veinte (20) imágenes adicionales por clase. Se obtuvieron los siguientes resultados:

- Durante el entrenamiento, el modelo convergió, aproximadamente, en la época n.º 30.
- Los valores de precisión y pérdida no se vieron afectados de forma negativa, es decir, no decayeron por más que el modelo había conseguido valores ideales. Estos valores son de  $\approx 0,1$  para la función de pérdida y 0,97 para el valor de precisión.
- En la fase de pruebas, se clasificó correctamente al 80 % de las imágenes pertenecientes al conjunto designado para pruebas.

En la segunda fase, se realizaron pruebas con especímenes de peces reales. Las clases escogidas fueron cojinova y bonito. Se trabajó con un espécimen por cada clase indicada. Los resultados obtenidos son los siguientes: respecto a la detección, se detectaron correctamente todos los peces presentados al sistema; respecto a la clasificación, el sistema clasificó correctamente el 80 % de las veces; sin embargo, en el 20 % de las pruebas realizadas con la clase cojinova se obtuvieron falsos positivos.

## 5.1 Discusión de resultados

Respecto a la primera fase de pruebas, se concluye que los resultados obtenidos durante la detección fueron satisfactorios porque alcanzaron el 90 % de exactitud. La solución planteada

para la detección de peces solo puede fallar si el pez tiene una forma distinta a la de un pez convencional o si el objeto a detectar no es un pez y tiene una forma similar a la de uno.

Respecto a los resultados de la tarea de clasificación, consideramos que son ideales. En términos de convergencia del modelo, se obtuvo convergencia en  $\approx$  al 80 % del proceso de entrenamiento. Esto se da porque el conjunto de datos es pequeño y, por la propia estructura y diseño del modelo, se consiguen valores ideales de manera rápida. Asimismo, la convergencia no se vio afectada por la continuidad de entrenamiento, en el resto de épocas, razón por la cual se muestra que el modelo es consistente.

Respecto a la precisión del modelo, plasmándola en términos sencillos, de cada diez imágenes capturadas de especímenes, ocho son clasificadas de manera correcta. Esto implica que los resultados obtenidos presentan un nivel elevado de precisión. De igual forma, realizando un análisis del porqué no es efectivo en su totalidad, concluimos que la inclusión de un número mayor de muestras representativas en el proceso de entrenamiento contribuiría a lograr el objetivo esperado.

Finalmente, de la segunda fase de pruebas se concluye que, respecto a la detección, se logró la detección satisfactoria de todos los peces reales en condiciones ideales de luminosidad. En lo que concierne a la tarea de clasificación, como se comentó, los falsos positivos que se obtienen son consecuencia de la similitud morfológica que tienen las clases estudiadas. El equipo de trabajo concluye que para optimizar este resultado es ideal contar con un conjunto de datos más extenso para el entrenamiento, de tal manera que el modelo de clasificación pueda generalizar mejor las clases de pescados incluidas.

## REFERENCIAS

- Aghdam, O. A. (s. f.). Deep Learning Features for Fish Classification, 3-5.
- Brown, M., y Lowe, D. G. (2002). Invariant Features from Interest Point Groups. *British Machine Vision Conference*. doi:10.1.1.1.8475
- Calvo Dopico, D. (2015). Implantación de la trazabilidad y su relación con la calidad: marco conceptual y retos estratégicos. Aplicación al sector pesquero. *Economía Agraria y Recursos Naturales*, 15(1), 79-98. doi:10.7201/earn.2015.01.05.a
- Caribbean Regional Fisheries Mechanism (CRFM). (2016). Manual de sistemas de trazabilidad del pescado y productos pesqueros. *Instituto Interamericano de Cooperación para la Agricultura, CRFM Publicación Especial*, (13).
- Chen, G., Sun, P., y Shang, Y. (2017). Automatic fish classification system using deep learning. *2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI)*, 24-29. doi:10.1109/ICTAI.2017.00016

- Chincha, R., y Tian, Y. (2011, November). Finding objects for blind people based on SURF features. *2011 IEEE International Conference on Bioinformatics and Biomedicine Workshops (BIBMW)*. (526-527). IEEE.
- Edelman, S., Intrator, N., y Poggio, T. (1997). Complex cells and object recognition. *Computer Science*.
- FAO. (2018). El estado mundial de la pesca y la acuicultura 2018. Cumplir los objetivos de desarrollo sostenible. Roma. Licencia: CC BY-NC-SA 3.0 IGO
- Fishing for a Future - Nature Works Everywhere. (s. f.). Recuperado de <https://www.natureworkseverywhere.org/resources/fishing-for-a-future/>
- Fouad, M. M., Zawbaa, H. M., El-Bendary, N., y Hassanien, A. E. (2013). Automatic Nile Tilapia fish classification approach using machine learning techniques. *13th International Conference on Hybrid Intelligent Systems (HIS 2013)*, 173-178. doi:10.1109/HIS.2013.6920477
- Frost, J., Geisler, T., y Mahajan, A. (s. f.). Monitoring illegal fishing through image classification.
- Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 36. Recuperado de <https://www.rctn.org/bruno/public/papers/Fukushima1980.pdf>
- Harris, C., y Stephens, M. (1988). A combined corner and edge detector. *Proceedings of the Alvey Vision Conference 1988*. doi:10.5244/C.2.23
- He, K., Gkioxari, G., Dollár, P., y Girshick, R. (2017). Mask R-CNN. *2017 IEEE International Conference on Computer Vision (ICCV)*, 2980-2988. doi:10.1109/ICCV.2017.322
- He, K., Zhang, X., Ren, S., y Sun, J. (2015). Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770-778. doi:10.1109/CVPR.2016.90
- Huang, G., Liu, Z., van der Maaten, L., y Weinberger, K. Q., (2017). Densely connected convolutional networks. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2261-2269. doi:10.1109/CVPR.2017.243
- Josifovski, J. (s. f.). Object recognition sift vs convolutional neural networks. Recuperado de [https://tams.informatik.uni-hamburg.de/lehre/2015ws/seminar/ir/pdf/slides/JosipJosifovskiObject\\_Recognition\\_SIFT\\_vs\\_Convolutional\\_Neural\\_Networks.pdf](https://tams.informatik.uni-hamburg.de/lehre/2015ws/seminar/ir/pdf/slides/JosipJosifovskiObject_Recognition_SIFT_vs_Convolutional_Neural_Networks.pdf)
- Juan, L., y Gwun, O. (2009). A comparison of SIFT, PCA-SIFT and SURF. *International Journal of Image Processing (IJIP)*, 3(4), 143-152.
- Karami, E., Prasad, S., y Shehata, M. (2015). Image matching using SIFT, SURF, BRIEF and ORB: Performance comparison for distorted images. *Newfoundland Electrical and Computer Engineering Conference*.

- Krizhevsky, A., Sutskever, I., y Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances In Neural Information Processing Systems*, 1-9. doi/10.5555/2999134.2999257
- Kruger, N., Janssen, P., Kalkan, S., Lappe, M., Leonardis, A., Piater, J., ... Wiskott, L. (2013). Deep hierarchies in the primate visual cortex: What can we learn for computer vision? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8), 1847-1871. doi:10.1109/TPAMI.2012.272
- LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., y Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Computation*, 1(4), 541-551. doi:10.1162/neco.1989.1.4.541
- Lowe, D. G. (1999). Object Recognition from Local Scale-Invariant Features (*SIFT*). *International Conference on Computer Vision*. doi:10.1109/ICCV.1999.790410
- Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2), 91-110. doi:10.1023/B:VISI.0000029664.99615.94
- Mahajan, A., y Perkins, A. Whale Detection and Identification from Aerial Photography.
- Matai, J., Kastner, R., Cutter, G. R., y Demer, D. A. (2012). Automated techniques for detection and recognition of fishes using computer vision algorithms. *NOAA Technical Memorandum*, 35-37.
- Meng, Y., y Tiddeman, B. (2006). Implementing the Scale Invariant Feature Transform (SIFT) Method. Citeseer.
- Nadarajan, G., Chen-Burger, J., Fisher, R., y Spampinato, C. (2011). A flexible system for automated composition of intelligent video analysis. *ISPA 2011-7th International Symposium on Image and Signal Processing and Analysis*, 259-264
- Piccardi, M. (2004). Background subtraction techniques: a review. *2004 IEEE International Conference on Systems, Man and Cybernetics*. doi:10.1109/ICSMC.2004.1400815
- Ren, S., He, K., Girshick, R., y Sun, J. (2016). Faster R-CNN: Towards real-time object detection with region proposal networks. 1-14. doi:10.1109/TPAMI.2016.2577031
- Shujuan, S., Zhize, X., Xingang, W., Guan, H., Wenqi, W., y De, X. (2015, May). Real-time vehicle detection using Haar-SURF mixed features and gentle AdaBoost classifier. *The 27th Chinese Control and Decision Conference (2015 CCDC)*, 1888-1894. IEEE.
- Simonyan, K., y Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

- Sobral, A., y Vacavant, A. (2014). A comprehensive review of background subtraction algorithms evaluated with synthetic and real videos. *Computer Vision and Image Understanding*, 122, 4-21. doi:10.1016/j.cviu.2013.12.005
- Sociedad Nacional de Pesquería (2018). Exportaciones pesqueras. Recuperado de <https://www.snp.org.pe/wp-content/uploads/2019/02/REPORTE-DE-LAS-EXPORT-PESQUERAS-A%C3%91O-2018.pdf>
- Suleiman, A., Chen, Y. H., Emer, J., y Sze, V. (2017). Towards closing the energy gap between HOG and CNN features for embedded vision (invited paper). *Proceedings - IEEE International Symposium on Circuits and Systems*. doi:10.1109/ISCAS.2017.8050341
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., y Rabinovich, A. (2014). Going deeper with convolutions. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1-9. doi:10.1109/CVPR.2015.7298594
- Tareen, S. A., y Saleem, Z. (2018). A comparative analysis of SIFT, SURF, KAZE, AKAZE, ORB, and BRISK. *2018 International Conference on Computing, Mathematics and Engineering Technologies (ICOMET)*, 1-10. doi:10.1109/ICOMET.2018.8346440
- Treiber, M. A. (2010). An introduction to object recognition: selected algorithms for a wide variety of applications. *Springer Science & Business Media*.
- Vedaldi, A. (2006). An implementation of SIFT detector and descriptor. University of California at Los Angeles, 7.
- Zeiler, M. D., y Fergus, R. (2014). LNCS 8689 - Visualizing and understanding convolutional networks. *Computer Vision, ECCV 2014 - 13th European Conference, Proceedings*. Recuperado de <https://cs.nyu.edu/~fergus/papers/zeilerECCV2014.pdf>



# Sistema para incremento de vocabulario para la comprensión lectora en primaria con ayuda de realidad aumentada

Gianella Urday  
gianellaurdayi@hotmail.com / Universidad de Lima, Perú

Recepción: 22-6-2019 / Aceptación: 8-8-2019

**RESUMEN.** El Perú presenta resultados insatisfactorios en evaluaciones internacionales y nacionales en el área de lectura. A medida que los grados avanzan, los porcentajes de satisfacción van disminuyendo. El problema radica en que los estudiantes no son capaces de comprender lo que están leyendo en su totalidad debido a que inician su vida académica con un bajo rendimiento. La presente investigación se enfoca en la baja comprensión lectora ocasionada por un vocabulario reducido para lo cual se desarrolló una aplicación en realidad aumentada (AR) para comprobar si esta tecnología aumenta el vocabulario y así mejorar la comprensión lectora. Los alumnos participantes fueron divididos en dos grupos, uno que no hace uso de la aplicación y otro que sí lo utiliza, de control y experimental, respectivamente. Ambos grupos realizaron una prueba basada en la Evaluación Censal de Estudiantes (ECE), evaluación nacional que da a conocer los niveles de aprendizaje en los que se encuentran los estudiantes peruanos, para posteriormente ser comparados. Los resultados mostraron una diferencia de 0,3, entre los puntajes promedios de ambos grupos, y de 0,8, en el grupo experimental. Se demuestra que el uso de la aplicación basada en realidad aumentada es útil para el aumento de vocabulario y, por lo tanto, beneficioso en el ámbito académico porque permite que los estudiantes realicen sus actividades de forma más dinámica y a la vez se concentran más, ayudándoles a entender los temas con mayor facilidad.

**PALABRAS CLAVE:** comprensión lectora, educación, realidad aumentada, vocabulario reducido

## A System to Improve Vocabulary for Enhancing Reading Comprehension in Primary School With the Help of Augmented Reality

**ABSTRACT.** Peru has achieved unsatisfactory results in international and national reading assessments. The higher the school grade, the less the satisfaction percentages. The problem is that students are not able to fully understand what they are reading because they start their academic life with a low reading comprehension performance. This research focuses on the low reading comprehension caused by a poor vocabulary, to which end an augmented reality-based application was developed to check if this technology improves the vocabulary and hence the reading comprehension. The participating students were divided into two groups: one that did not use the application and the other one that used it—i.e., a control group and an experimental group, respectively. Both groups took a test based on the Student Census Assessment (ECE), a national assessment that discloses the learning levels in which Peruvian students are, and those tests were compared later. The results showed a difference of 0.3 between both groups' average scores, and a difference of 0.8 in the experimental group. The research demonstrates that this augmented reality-based application is useful to improve the vocabulary and, therefore, is beneficial in the academic field, since it allows students to carry out activities more dynamically and at the same time concentrate more, helping them to understand the topics more easily.

**KEYWORDS:** reading comprehension, education, augmented reality, poor vocabulary



## 1. INTRODUCCIÓN

La evaluación internacional PISA tiene como objetivo evaluar los conocimientos y habilidades de los estudiantes de quince años entre los países participantes. Los resultados se categorizan en ocho niveles: debajo de nivel 1b, nivel 1b, nivel 1a, nivel 2, nivel 3, nivel 4, nivel 5 y nivel 6. PISA señala que el nivel 2 es el mínimo requerido para ser participantes de la sociedad actual (OECD, 2016, p. 5).

De acuerdo con la evaluación PISA 2015 en el área de lectura, los resultados de los estudiantes peruanos de quince años están por debajo de la media total con una puntuación de 398, siendo 535 la más alta y 347 la más baja, de Singapur y Líbano respectivamente. En el caso del Perú, solo un 82,4 % de la población de quince años está matriculada en secundaria, en donde los niveles de desempeño predominantes son el nivel 1a y el nivel 2 con 28,3 % y 27,3 % respectivamente (Minedu, 2016a).

Por otro lado, la Evaluación Censal de Estudiantes (ECE), realizada por el Ministerio de Educación del Perú, sirve para dar a conocer los niveles de aprendizaje de los alumnos tanto de escuelas públicas como privadas. De esta evaluación los datos relevantes para la investigación son los resultados en el área de lectura en los grados de segundo y cuarto de primaria, y segundo de secundaria. Los resultados se categorizan en cuatro niveles: previo al inicio (solo para cuarto de primaria y segundo de secundaria), en inicio, en proceso y satisfactorio (Minedu, 2016b, p. 3).

Los niveles 5 y 6 de la evaluación PISA equivalen al nivel de logro satisfactorio del ECE en donde los estudiantes han sido capaces de cumplir con las definiciones de competencia lectora de cada evaluación.

El problema radica en que los porcentajes de satisfacción del ECE van disminuyendo a medida que los grados avanzan. Pasa de un 46,4 % a un 31,4 % y de este a un 14,3 % en los grados de segundo y cuarto de primaria, y segundo de secundaria, respectivamente. Para ser considerado satisfactorio el nivel del estudiante, debe tener “la capacidad para comprender, utilizar, reflexionar e interesarse por los textos escritos, para alcanzar los propios objetivos, desarrollar el conocimiento y potencial personales, y participar en la sociedad” (UMC, 2016, p. 43).

Braslavsky (2005), Bravo, Villalón y Orellano (2005) definen a la lectura como una actividad que involucra procesos cognitivos y psicolingüísticos complejos, la decodificación y la comprensión. Vallés Arándiga (2005) la define como la interpretación de la letra para darle un significado y a partir de ella se elabora una comprensión de texto (p. 49). Si el estudiante no es capaz de desarrollar la lectura, su crecimiento como persona, al igual que el desarrollo de aptitudes como el pensamiento crítico, la reflexión o la sensibilidad se verán afectados (Carvajal, 2015, p. 297).

Existe la inquietud de que los alumnos que empiezan con bajo rendimiento en el área de lectura lo seguirán teniendo a lo largo de su educación, sea colegio o universidad (Bravo *et al.*, 2005).

Además, por ser una actividad que se realiza más por imposición que por interés. La comprensión lectora es parte fundamental para una educación exitosa, ya que se hace uso de ella en diferentes materias. Por ejemplo: al resolver un problema matemático es indispensable comprender lo leído en el enunciado, también para realizar una mezcla de químicos deben leerse las instrucciones. Con el avance de los grados, la dificultad de los enunciados, las instrucciones, las preguntas va en aumento. Las personas que se mantienen con una baja comprensión lectora tienen un menor desarrollo de sus capacidades a comparación de las demás (Minedu, 2016b, p. 52).

Los conocimientos que se adquieren durante los años académicos permiten que la persona desarrolle las capacidades necesarias para el profesional que debe ser, pues una de las formas por las que se adquiere información es mediante la lectura. Lo importante de ella es comprender lo que uno está leyendo, para que así la persona forme sus propias ideas y opiniones.

Dentro de los factores que ocasionan la baja comprensión lectora está el vocabulario reducido. Para Defior (1996), el alumno va a tener la dificultad de entender lo que quiere transmitir el texto si es que no tiene claro lo que significa (p. 4). Como el alumno no reconoce la palabra, se entiende que no tiene información almacenada sobre ella, por lo cual al momento de leer no será capaz de relacionar su conocimiento previo con el nuevo que la lectura le brinda (p. 3).

Canet-Juric, Burin, Andrés y Urquijo (2013) presentan el perfil cognitivo de niños con rendimientos bajos en comprensión lectora: mayor dificultad para almacenar información durante períodos cortos de tiempo, dificultad de prestar atención, incapacidad de mantener la relación entre las oraciones que se van leyendo, bajo rendimiento en la habilidad de vocabulario y definición de palabras (p. 1003).

El principal objetivo de esta investigación es desarrollar una aplicación basada en realidad aumentada para el incremento del vocabulario que contribuya a la comprensión lectora.

## 2. ESTADO DEL ARTE

### 2.1 Comprensión lectora

Vallés Arándiga y Vallés Tortosa (2006) presentan herramientas que son eficaces para mejorar la comprensión lectora: identificar la idea principal del texto, realizar organizadores gráficos, usar palabras suplentes, elaborar preguntas y respuestas sobre el texto, expresar la lectura con otras palabras, elaborar resúmenes.

La profesora Salazar (2018) menciona tres niveles de comprensión lectora que se trabajan con los alumnos de educación primaria: nivel literal, nivel inferencial y nivel crítico valorativo. En el segundo grado de primaria se desarrollan los dos primeros niveles, con la ayuda de diferente tipología textual. El enfoque utilizado es el comunicativo textual, en donde todo el trabajo a realizar es aprender nuevas palabras, nuevos tipos de textos, gramática, como parte de

un texto (Salazar, comunicación personal, 15 de mayo del 2018). Los niveles de comprensión lectora que la profesora Correa menciona equivalen al nivel de comprensión base del texto 1: comprensión léxica, sintáctica y referencial local de Fernández (2004). Para el nivel base del texto 1, las pruebas para evaluar la comprensión lectora son de tipo objetivas, semiobjetivas y de pensamiento en voz alta (p. 425). Gutiérrez-Braojos y Salmerón Pérez (2012) concuerda con Fernández (2004) en el uso de la técnica del pensamiento en voz alta y añade el uso de cuestionarios.

Soriano, Vidal-Abarca y Miranda (1996) compararon dos procedimientos de instrucción en comprensión y aprendizaje de textos, instrucción directa y enseñanza recíproca. En el estudio participaron veintiún niños del quinto grado de primaria designados en tres grupos de siete alumnos cada uno, el primero educado con la instrucción directa, el segundo con la enseñanza recíproca, ambos por doce sesiones, y el último, grupo de control, por ninguno. Se demostró que los grupos quienes aplicaron los métodos de aprendizaje tuvieron un mayor puntaje que el último grupo en la parte de captación de ideas principales y resumen, y que el segundo grupo mejoró en la conciencia de estructura textual (p. 59).

## 2.2 Aprendizaje social

Schunk (2012) indica que una considerable parte del aprendizaje proviene de observar y escuchar ejemplos, acciones o conductas, y que el aprendizaje se realiza de manera más rápida en comparación a que si el sujeto realizara lo anteriormente mencionado. Esto evita que se experimenten situaciones que podrían traer consecuencias negativas. El uso de este aprendizaje permite a los profesores retroalimentar de forma correctiva para afinar las destrezas y habilidades de los estudiantes (p. 121).

Rebollo Catalán, García Pérez, Buzón García y Barragán Sánchez (2012) muestran resultados positivos tras la implementación del aprendizaje colaborativo en comunidades virtuales. De los 113 estudiantes que participaron, el 62,7 % marcó la experiencia como positiva y el 22 % como muy positiva. La valoración de la metodología usada se encuentra en un rango de 1 a 7, como resultado los alumnos otorgaron una puntuación media global de 5,1 (p. 116). Este resultado se basa en que los usuarios valoraron la oportunidad de poder colaborar con el aprendizaje de los demás, compartiendo conocimientos entre sí (p. 117).

Barba Martín (2010) compara el aprendizaje cooperativo y la asignación de tareas para confirmar la existencia de alguna diferencia de importancia en su uso en la enseñanza del *acrosport* (p. 14). Los resultados enfocados en la calidad de trabajo muestran que el aprendizaje cooperativo tiene un mejor resultado si se añade la creatividad de la realización de la tarea a su eficacia de 1,4 puntos, dentro de una escala entre 1 y 5, y muestran mayor seguridad en la ejecución de la parte creativa (p. 17).

Camilli Trujillo, López Gómez y Barceló Cerdá (2012) indica al método de divisiones del trabajo de los equipos de estudiantes (Student Teams Achievement Divisions, STAD) como el más eficiente y al método Jigsaw como el menos eficiente, debido a que no incorpora la responsabilidad individual como recompensa en el grupo cooperativo (p. 92). El rendimiento de los alumnos no es afectado de forma significativa si no hay una recompensa grupal, más bien, la responsabilidad individual es de suma importancia ya que los alumnos demuestran su compromiso y esfuerzo en las tareas que se les asignan para poder lograr el mejor resultado entre todos (p. 95).

### 2.3 Hipertexto

Rueda (2001) señala que los estudios sobre el uso pedagógico de hipertextos se enfocan en el aprendizaje asociado con los conocimientos previos, estilos y estrategias cognitivas, el uso de la tecnología, entre otros (p. 279).

Los estudiantes que tienen conocimientos previos sobre un tema y utilizan el hipertexto para estudiar, haciendo uso de buenas estrategias cognitivas, muestran un mejor desempeño en su aprendizaje. Lo contrario sucede con los estudiantes principiantes en un tema, que no entienden del manejo de los hipertextos y tienen una baja capacidad de estudio, además les genera disgusto el hacer uso de hipertextos (p. 280).

Martínez, Astiz, Medina, Montero y Pedrosa (1998) probaron dos formas en las cuales se hace uso del hipertexto, los llamó “árbol” y “libro”, que constan de 172 nodos y 91 páginas, respectivamente. El uso del método del libro fue beneficioso para los usuarios que no tenían experiencia con el uso del hipertexto; sin embargo, el tiempo que se tomaron para desarrollar las indicaciones fue mayor debido a la cantidad de pantallas que recorrían. Los estudiantes que hicieron uso del método del árbol presentaron dificultad en su manejo por lo que en algunos casos tuvieron que recurrir a guías. Ninguno de los dos influyó negativamente en el desarrollo de las tareas que se les indicaba a los estudiantes por lo que se demostró que son eficaces.

### 2.4 Realidad aumentada

Martín-Sabarís y Brossy-Scaringi (2017) realizaron un estudio exploratorio sobre la realidad aumentada aplicada al aprendizaje en personas con síndrome de Down. Su uso proporciona herramientas audiovisuales que permiten a los usuarios recordar con mayor facilidad ideas y conceptos debido al impacto emocional que les produce (p. 747).

Construct 3D es una aplicación que posibilita la creación de nuevos diseños con figuras geométricas de forma colaborativa. Ha sido probada en el Museo de Tecnología de Viena, facilitando el aprendizaje de cursos relacionados con matemáticas y geometría (Kaufmann y Schmalstieg, 2002).

El diario *El País* implementó un servicio en donde los lectores escanean un código situado encima de la información. Esta funcionalidad fue implementada para brindar información actualizada sobre las elecciones presidenciales francesas de 2017 (Parra Valcarce, Edo Bolós y Marcos Recio, 2017, p. 1681).

Barba Vera, Yasaca Pucuna y Manosalvas Vaca (2015) muestran los resultados conseguidos en su investigación sobre la mejora del aprendizaje de los estudiantes de medicina en el curso de Anatomía III con ayuda de realidad aumentada. Se obtuvo una diferencia de  $P$  valor = 0,012 entre el grupo que hizo uso de la herramienta con el que no. El primer grupo tuvo una media de 3,759 y desviación estándar de 0,711. El segundo grupo, una media de 3,484 y desviación estándar de 0,757. En conclusión, el uso de la realidad aumentada contribuyó de forma positiva en el aprendizaje (p. 1428).

### 3. MARCO TEÓRICO

#### 3.1 Realidad aumentada

La realidad aumentada (RA) como concepto aparece en 1990, Hedley, Billingham, Postner, May y Kato (2002) y Milgram *et al.* (1994) concuerdan que es “una variación de los entornos virtuales” (p. 739). A diferencia de la realidad virtual, en donde el usuario está inmerso en un mundo ficticio, pese a dirigir objetos y contextos de apariencia real (Parra Valcarce *et al.*, 2017, p. 1672), la realidad aumentada permite que los objetos virtuales coexistan en la realidad. Prendes (2015) aclara que la realidad aumentada no tiene como propósito reemplazar a la física, sino permitir el contacto con lo virtual en la realidad física (p. 71).

Para Azuma (2017) la realidad aumentada tiene que desarrollar tres características para ser llamada como tal, la combinación de objetos reales y virtuales, que sea en tiempo real y en 3D (p. 740).

De acuerdo con la Fundación Telefónica (2011), la realidad aumentada necesita cuatro elementos fundamentales: un dispositivo por el cual se pueda visualizar el entorno real, un soporte físico sobre el cual se mostrará el resultado, un procesador que pueda interpretar los datos del mundo real y que muestre el contenido virtual correspondiente y un elemento que desencadene la realidad aumentada (p. 11).

Existen diversos modelos de realidad aumentada según las tecnologías desarrolladas, soportes, niveles y diversos criterios (Cabero Almenara, Leiva Olivencia, Moreno Martínez, Barroso Osuna y López Meneses, 2016, p. 77).

### 3.1.1 Tecnologías desarrolladas

Dentro de las tecnologías desarrolladas están los patrones de disparo de *software* RA, como marcadores y códigos QR. Los marcadores necesitan tener un patrón único para que la cámara examine y determine el objeto que se debe mostrar (Nava, González, Galicia y Flores, 2016, p. 5). Puede tomar cualquier forma, a diferencia de los códigos QR en donde siempre tiene que haber tres cuadrados en las esquinas.

### 3.1.2 Soporte

El soporte utilizado es un dispositivo móvil, se requiere el uso de un *smartphone* o *tablet* con cualquier sistema operativo y que tenga cámara digital. Ruiz Torres (2013, p. 84) agrega que estos dispositivos agregan las imágenes o gráficos virtuales antes de que se exhiban en la pantalla (p. 32). Su variedad de tamaños los hacen menos intrusivos, más accesibles y fáciles de manejar.

### 3.1.3 Niveles

De acuerdo con Reinoso (2016), el nivel de la realidad aumentada es según la complejidad del sistema. Para este caso, el nivel es el número 1 debido a que está basado en marcadores (p. 740).

Bacca, Baldiris, Fabregat, Graf y Kinshuk (2014) concluyen sobre la realidad aumentada: la intención principal de usar la RA es para explicar temas de formas interactivas. Además, los experimentos sobre la RA en juegos educativos siguen en proceso, su uso promueve la motivación, interacción y colaboración en el aprendizaje y es capaz de mantener interesados a los alumnos en el tema (p. 146).

## 3.2 Evaluaciones

La Evaluación Censal de Estudiantes (ECE) presenta en un informe para docentes el tipo de preguntas que los maestros deben evaluar a los alumnos respecto a la comprensión de textos y un estándar de respuestas que los estudiantes deberían dar. Se presenta un modelo de evaluación de lectura compuesto por tres elementos: capacidades, textos y contextos de lectura (p. 6).

Tabla 1  
Elementos para la evaluación de la competencia de lectura

Elementos	Descripción
Capacidades	Identificar información de forma explícita, inferir el significado del texto y entender la lectura en partes específicas y como conjunto
Textos	Narrativo-descriptivo de forma continua
Contextos de lectura	Educacional

Elaboración propia

Los resultados de estas evaluaciones se distribuyen en tres niveles: satisfactorio, en proceso y en inicio (Minedu, 2015a, p. 4). Para que el alumno logre un nivel satisfactorio, es decir, que sí logró los aprendizajes esperados, debe cumplir con las capacidades previamente mencionadas.

Estos elementos se pueden visualizar en las evaluaciones del cuadernillo *Demostrando que aprendimos* para el segundo grado de primaria.

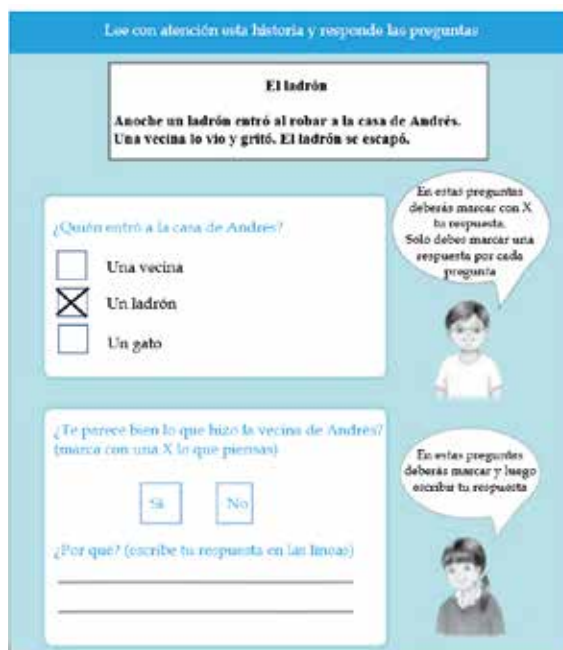


Figura 1. Preguntas ejemplo en base al texto mostrado

Fuente: Ministerio de Educación del Perú

La figura 1 muestra preguntas ejemplo basadas en la guía metodológica realizada por la UNICEF. La primera pregunta es de opción múltiple con única respuesta que se encuentra de forma explícita en el texto. La segunda pregunta abarca el nivel crítico-valorativo de los estudiantes, preguntando sobre el reconocimiento de las buenas y malas acciones. También se trabaja el nivel inferencial al pedir la opinión sobre lo sucedido.

## 4. METODOLOGÍA

### 4.1 Etapas de solución

El plan de desarrollo de la propuesta de solución se plantea a continuación y se resume en la figura 2.

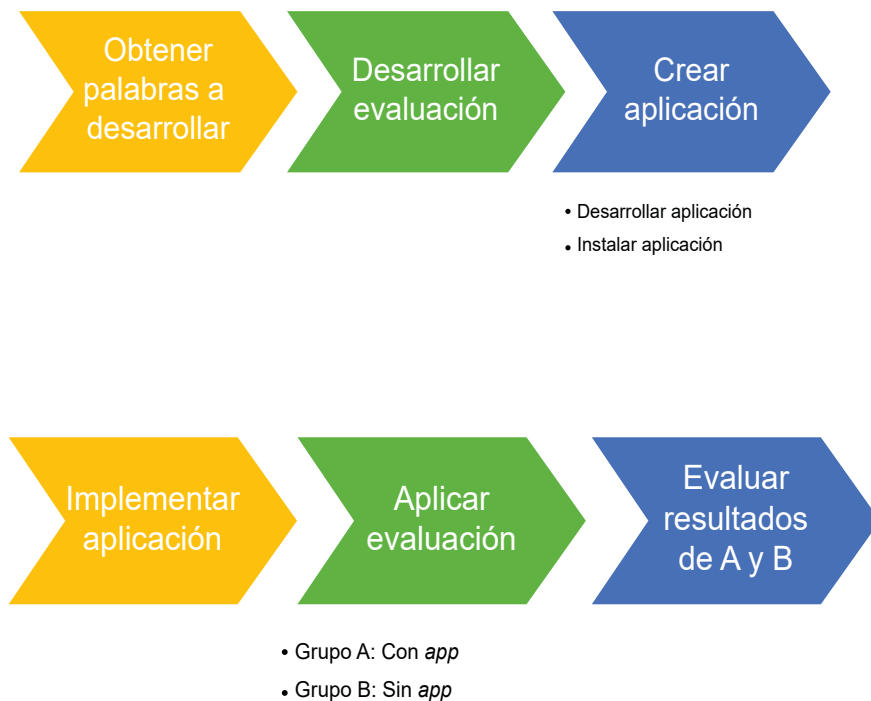


Figura 2. Plan de desarrollo

Elaboración propia

*Primera etapa.* Obtener palabras a desarrollar: se realiza la lectura del plan lector, para después elegir un capítulo en específico. Las palabras escogidas serán las que la docente



encargada del grupo a evaluar ha decidido utilizar como parte de sus evaluaciones de comprensión lectora.

*Segunda etapa.* Desarrollar evaluación: las palabras seleccionadas del capítulo del plan lector serán utilizadas en la evaluación, relacionándolas con preguntas sobre la comprensión de lectura, para validar la efectividad de la realidad aumentada en la mejora del vocabulario para la comprensión lectora.

*Tercera etapa.* Crear aplicación: el aplicativo se realiza en el programa Unity conjunto a Vuforia. Luego se crean los sustantivos en 3D para poder importarlos en el programa Unity. Una vez terminado, se instalará en los dispositivos.

*Cuarta etapa.* Implementar aplicación: después de que los alumnos hayan terminado de leer el capítulo respectivo, se les proporcionarán los dispositivos con la aplicación instalada. Ellos harán uso de los dispositivos para obtener nueva información sobre el vocabulario de la lectura.

*Quinta etapa.* Aplicar evaluación: una vez que los alumnos hayan terminado de usar la aplicación, se pasa a la evaluación respecto al capítulo leído. Serán divididos en dos grupos, uno que no hayan utilizado la aplicación y otro en la que sí.

*Sexta etapa.* Analizar y comparar resultados: finalmente, se compararán los resultados de ambos grupos para examinar el impacto del uso de la realidad aumentada en el aprendizaje de vocabulario. Como indicadores de que la implementación fue exitosa, tendremos el porcentaje de respuestas correctas por alumno en cada capítulo, porcentaje de respuestas correctas por capítulo y el puntaje promedio del grupo de control y experimental.

## 4.2 Métodos de investigación

Los métodos a utilizar en el presente trabajo son: implementación y experimentación. Según Berndtsson, Hansson, Olsson y Lundell (2008) el objetivo principal del método de implementación es demostrar que la solución actúa de determinada forma bajo ciertas eventualidades. Es primordial el uso de un buen *software* para asegurar la validez y confiabilidad del trabajo. La validez es importante ya que la implementación muestra la manera en la que la solución propuesta funciona. Con confiabilidad se refiere a que la solución resulta ser útil (p. 64).

Puesto que la investigación involucra una hipótesis, se hace uso del método de experimentación. Según Sampieri *et al.* (1998) “la esencia de esta concepción de experimento es que requiere la manipulación intencional de una acción para analizar sus posibles resultados” (p. 129). Los resultados de la experimentación son utilizados para comprobar o desmentir la hipótesis planteada (Berndtsson *et al.*, 2008, p. 65).

También se puede incluir el método de observación, debido a que a partir de las observaciones que se realicen a los sujetos de muestra, al momento de utilizar la solución, se

anotan ideas sobre lo que se puede hacer en adelante para la investigación (Evertson y Green, 1989, p. 176).

### 4.3 Tipo de investigación

De acuerdo con Hernández Sampieri, Fernández Collado y Baptista Lucio (2014), la investigación es de tipo experimental pura en la cual se realiza: modificación voluntaria de variables independientes, evaluación de variables dependientes, control y autenticidad, y cotejar entre dos o más grupos (p. 127).

### 4.4 Métodos de validación

Para la presente investigación la validación es dirigida a la metodología de la solución; es decir, validar que el procedimiento de la solución se esté realizando de manera correcta para cumplir con los objetivos planteados.

De acuerdo con la guía metodológica realizada por Unicef (2003), se utilizaron dos tipos de validación: una técnica y otra con población. Para la validación técnica se hizo una revisión de la solución por parte de especialistas, quienes se encargaron de cotejar que la intención de la propuesta de solución haya sido la adecuada para los alumnos de segundo grado de primaria. En el caso de la validación con población, implicaron entrevistas personales con preguntas sobre la solución relacionadas a los siguientes componentes: atracción, comprensión, involucramiento, aceptación e inducción a la acción.

La atracción se relaciona con la atención que el usuario le presta al material, si le atrajo lo visual o el contenido de él. La comprensión es sobre si el usuario entendió el contenido del material, para la propuesta sería si logró comprender el significado de la palabra. El involucramiento examina como el usuario se siente respecto al material, si cree que es el adecuado para él. La aceptación se refiere a que el usuario utiliza el material de forma asertiva, no muestra disconformidad. Finalmente, en la inducción a la acción se busca percibir si el usuario ha tenido algún cambio sea conductual o intelectual después del uso del material (Unicef, 2003).

## 5. EXPERIMENTACIÓN

### 5.1 Dispositivo a utilizar

El dispositivo en donde fue instalada la solución fue un Iphone 6s de 128 gb con la versión 12.3.1 del *software*, el tamaño de la pantalla es de 12 cm o 4,7 pulgadas. La cámara trasera es de 12 MP y la frontal de 5 MP.

## 5.2 *Image target*

La solución se desarrolló con el programa Unity junto al SDK Vuforia. La aplicación detecta palabras previamente acordadas para mostrar una imagen en 3D al estas ser reconocidas.

Vuforia requiere una imagen denominada *image target* (imagen objetivo) para que pueda mostrar otra imagen al momento de pasar el dispositivo encima de ella.

Como primera opción para *image target* se tuvo la página 9 completa del plan lector, *De la A a la Z Perú*, se colocó sobre la palabra *felino* el resultado que se quería mostrar. Entre las complicaciones, la más frecuente fue que el dispositivo en donde se instaló la aplicación tenía que estar a una distancia en donde la cámara pudiera captar toda la página y todas las palabras que se mostraran en 3D y aparecieran al mismo tiempo, creando confusión y distracción.



Figura 3. Resultado parcial

Fuente: *Plan lector De la A a la Z Perú*

La siguiente opción fue tener a la palabra como *image target*; sin embargo, al momento de subirla a la base de datos de Vuforia la calificación que se le daba a la imagen respecto a su capacidad de ser detectada variaba, para palabras de fuente de color claro era nula la capacidad de detección. Además, al momento de probar la aplicación no reconocía la palabra, puesto que habían otras antes y después de la seleccionada.



Figura 4. Palabra "felinos"

Fuente: *Plan lector De la A a la Z Perú*



Figura 5. Palabra “cajón” sin capacidad de ser detectada

Fuente: *Plan lector De la A a la Z Perú*

Como tercera opción se seleccionó todo un párrafo como *image target*. El resultado fue parcialmente exitoso, ya que a veces el programa demoraba en reconocer el párrafo como el *image target* subido a la base de datos de Vuforia. Además, la imagen resultante desaparecía de vez en cuando y se dejaba de reconocer al párrafo.

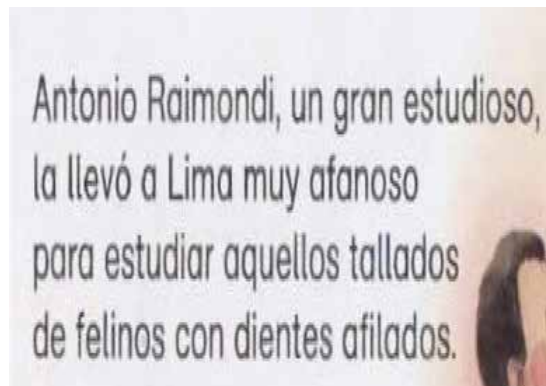


Figura 6. Párrafo seleccionado

Fuente: *Plan lector De la A a la Z Perú*

La cuarta opción fue el uso de QR o marcadores, aprovechando el espacio que hay entre los párrafos y los lados de las páginas. Para decidir cuál iba a ser utilizado, se les preguntó a los alumnos, que hicieron uso de la aplicación, cuál les gustaba y creían que era el más adecuado para la aplicación. El 72,73 % eligió el marcador personalizado, ya que lo veían más interesante y los ayudaba a relacionar con las palabras.

Fueron seleccionadas de cada párrafo entre una y dos palabras para ser mostradas por la aplicación. Se probó con diferentes tamaños de marcadores, 0,5 cm, 1 cm, 1,5 cm, 2 cm, 2,5 cm, 3,5 cm, 4 cm y 5 cm.

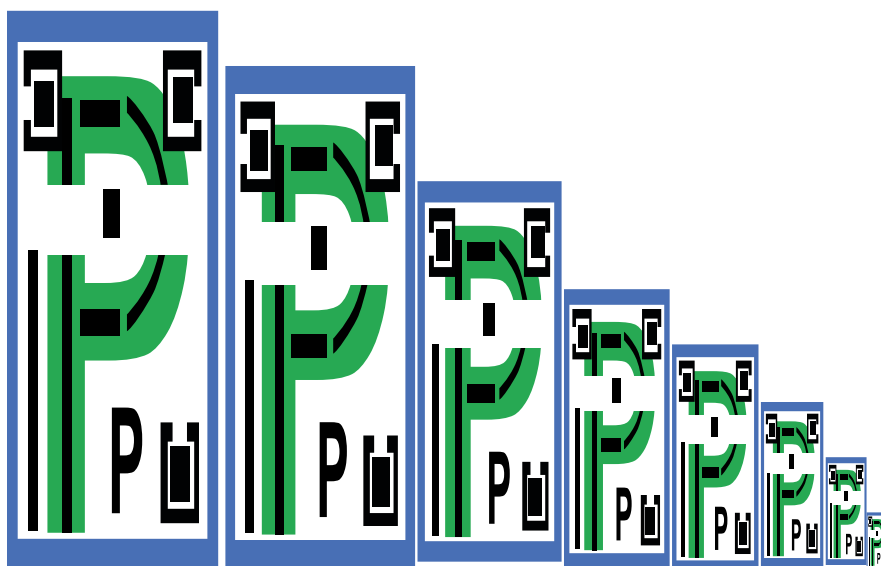


Figura 7. Diferentes tamaños de los marcadores  
Elaboración propia

Se realizaron comparaciones entre los tamaños de los marcadores respecto a la distancia mínima en la cual el *image target* era reconocido.

Tabla 2

Resultados de reconocimiento de QR de diferentes tamaños en relación con la distancia

Distancia/Tamaño	0,5 cm	1 cm	1,5 cm	2 cm	2,5 cm	3,5 cm	4 cm	5 cm
5 cm	NO	NO	NO	SÍ	NO	SÍ	SÍ	SÍ
10 cm	NO	NO	SÍ	SÍ	SÍ	SÍ	SÍ	SÍ
15 cm	NO	NO	SÍ	SÍ	SÍ	SÍ	SÍ	SÍ
20 cm	NO	NO	SÍ	SÍ	SÍ	SÍ	SÍ	SÍ
25 cm	NO	NO	NO	NO	SÍ	SÍ	SÍ	SÍ
30 cm	NO	NO	NO	NO	SÍ	SÍ	SÍ	SÍ

Elaboración propia

Tabla 3  
Resultados de marcadores de diferentes tamaños relacionados a si son distractivos o no

Distancia/Tamaño	0,5 cm	1 cm	1,5 cm	2 cm	2,5 cm	3,5 cm	4 cm	5 cm
Sí						X	X	X
No	X	X	X	X	X			

Elaboración propia

La elección del tamaño del marcador se basó en la mayor capacidad de reconocimiento de la *image target* con la menor distracción posible, es decir, que el tamaño del marcador sea el mínimo posible. Además, se tomó en cuenta la opinión del grupo experimental, el 27,27 % eligió el del 1 cm y el 18,18 %, el de 2 cm y 2,5 cm cada uno. Se seleccionó el tamaño de 2 cm. Las palabras que los usuarios van a poder visualizar en 3D se subrayan para poder facilitar su ubicación.

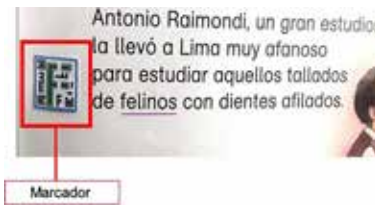


Figura 8. Palabra con QR personalizado  
Fuente: Plan lector De la A a la Z Perú

Se muestra en la figura 8 el marcador personalizado de la palabra “felinos”, el marcador tiene en el centro la primera letra de dicha palabra para que se relacione con mayor facilidad.

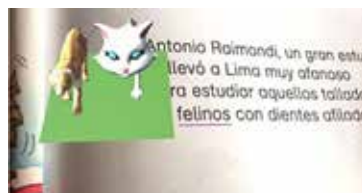


Figura 9. Resultado final  
Elaboración propia

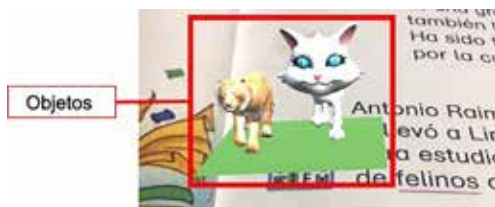


Figura 10. Otro ángulo del resultado final  
Elaboración propia



Figura 11. Muestra de dos resultados al mismo tiempo  
Elaboración propia

El resultado final consiste en la muestra en 3D de la palabra subrayada sobre una base para diferenciarlo del QR. Está inclinado de forma que el usuario es capaz de visualizarlo mejor. La solución también es capaz de mostrar múltiples resultados.

## 6. DISCUSIÓN DE RESULTADOS

La muestra consiste en veintidós alumnos de segundo grado de primaria, divididos en dos grupos: uno de control, con once alumnos, quienes no harían uso de la herramienta; el otro de experimentación, también con once alumnos.

### 6.1 Resultados del grupo de control

Tabla 4  
*Porcentajes de respuestas correctas de los capítulos por cada alumno sin app*

Capítulo / Alumno	1	2	3	4	5	6	7	8	9	10	11
A	33,3	83,3	50	50	50	100	66,7	50	50	50	56,7
E	66,7	0	33,3	66,7	83,3	50	33,3	50	33,3	50	40
F	75	87,5	37,5	75	75	75	37,5	50	100	62,5	60

Elaboración propia

La tabla 4 representa el porcentaje de respuestas correctas por alumno en cada capítulo. El alumno 9 obtuvo el porcentaje mayor con 100 % en el capítulo F, mientras que el alumno 2 obtuvo el porcentaje menor con 0 % en el capítulo E.

Tabla 5  
*Porcentaje de respuestas correctas por capítulo sin app*

Capítulo	Respuestas correctas
A	56,06
E	43,93
F	61,36

Elaboración propia

La tabla 5 representa el porcentaje de respuestas correctas por capítulo. El capítulo F es el que tiene el mayor porcentaje de respuestas correctas. Esto se debe a que la mayoría de alumnos entendió de forma exitosa el tema del texto y relacionaba con los conceptos que este mostraba. Al preguntarles sobre los conceptos, pudieron acordarse sobre lo leído y se ayudaron de sus conocimientos previos.



## 6.2 Resultados del grupo experimental

Tabla 6

*Porcentajes de respuestas correctas de los capítulos por cada alumno con app*

Capítulo / Alumno	1	2	3	4	5	6	7	8	9	10	11
A	83,3	50	83,3	66,7	83,3	83,3	50	50	66,7	50	66,7
E	33,3	16,7	33,3	33,3	66,7	50	50	16,7	33,3	100	66,7
F	37,5	37,5	37,5	75	37,5	62,5	37,5	62,5	87,5	75	87,5

Elaboración propia

La tabla 6 muestra los porcentajes de respuestas correctas de cada alumno en los tres capítulos. El alumno 10 obtuvo el porcentaje mayor con 100 % en el capítulo E, mientras que los alumnos 2 y 8 obtuvieron el porcentaje menor con 16,7 % en el mismo capítulo.

Tabla 7

*Porcentaje de respuestas correctas por capítulo con app*

Capítulo	Respuestas correctas
A	66,6
E	45,45
F	60,23

Elaboración propia

La tabla 7 representa el porcentaje de respuestas correctas por capítulo. Para los alumnos que usaron la aplicación, el capítulo con el número mayor de respuestas correctas fue el A. Varios de ellos complementaron sus respuestas a partir del objeto 3D que se mostraba para tal capítulo y lo relacionaban con conceptos previos.

Tabla 8  
*Comparación de puntaje del grupo de experimentación*

Alumno	Antes de <i>app</i>	Después de <i>app</i>
1	3,5	5
2	3	3,5
3	3,5	5
4	5	6
5	6,5	7
6	6	6,5
7	3,5	4,5
8	4,5	4,5
9	6	6,5
10	7	7,5
11	6,5	7,5

Elaboración propia

La tabla 8 muestra los puntajes obtenidos por el grupo de experimentación al hacer uso de la aplicación, el puntaje máximo a obtener es de 10 puntos.

La fila “Antes de *app*” contiene los puntajes que los alumnos del grupo de experimentación hubieran obtenido si no utilizaban la aplicación. Se observa un aumento de puntaje ya que el uso de la aplicación permitió que los alumnos fuesen capaces de captar la nueva información y complementar sus conocimientos previos.

### 6.3 Comparación de resultados

Tabla 9  
*Puntaje promedio del grupo de control y experimental*

Promedio grupo de control	Promedio grupo experimental
5	5,8

Elaboración propia

La diferencia de los promedios de ambos grupos es de 0,3 puntos.

Tabla 10  
*Puntaje promedio del grupo experimental al utilizar y no utilizar la app*

Experimental sin app	Experimental con app
5	5,8

Elaboración propia

Sin embargo, la diferencia de los promedios del grupo experimental si no hubieran utilizado y al hacer uso de la aplicación es de 0,8 puntos.

## 6.4 Validación

### 6.4.1 Validación por población

Tabla 11  
*Resultados de validación poblacional*

Componente	Pregunta	Respuesta promedio
Atracción	¿Qué es lo que más le gusta? ¿Por qué?	Poder ver los objetos en 3D porque son interesantes y llamativos.
	¿Qué es lo que menos le gusta? ¿Por qué?	Nada.
	¿Piensa que a otras personas de su edad o grado les guste? ¿Por qué?	Sí, por el 3D.
Comprensión	¿De qué se trata la solución explicada?	Usar un celular para ver dibujos en 3D.
	¿Qué ha resultado difícil de la explicación?	El porqué de la palabra subrayada y su relación con el QR.
	¿Entiende cómo funciona la aplicación?	Sí entienden.
Involucramiento	¿La solución es dirigida para personas de su edad o grado?	Sí.
	¿Hay algo que resulte familiar?	No.
Inducción a la acción	¿Algo resulta incómodo o molesto?	Nada.
	La aplicación requiere que se haga algo ¿Qué es?	Cuando se encuentre con la palabra subrayada se utiliza el celular para ponerlo encima del cuadro y ver el dibujo en 3D.
	¿Cree que necesitaría ayuda para utilizarla?	Sí, de un adulto.

Elaboración propia

La tabla 11 muestra las respuestas promedio respecto a las preguntas realizadas en base a las presentadas para la validación poblacional de la guía metodológica de Unicef (2003). El 72,72 % de los alumnos participantes les pareció muy atractiva la idea de poder ver objetos en 3D mientras leían. Además, les daba curiosidad la existencia de una palabra subrayada en el texto. El 63,64 % de los alumnos opinan que necesitarían ayuda, sea de su profesora o algún adulto a cargo, para utilizar la aplicación de la forma correcta. Al 45 % aproximadamente de los alumnos se les volvió a explicar la solución debido a que había conceptos con los que no estaban familiarizados, como “marcadores”, “realidad aumentada” y la relación entre los marcadores con las palabras subrayadas.

### 6.4.2 Validación por expertos

Tabla 12

*Resultados de validación técnica*

Pregunta	Respuesta
Comentarios sobre la solución	El tema de los dibujos es llamativo. La ayuda virtual ayuda a que la actividad sea más significativa.
Comentarios sobre el contenido que se muestra	El contenido es adecuado para los niños, ya que a los chicos les gusta mucho los temas tridimensionales.
¿La solución mostrada es adecuada para los alumnos del segundo grado de primaria?	Sí, el hecho de ver los objetos en 3D ayuda con la motivación de los niños para que lean.
¿Cree que se puede mejorar? ¿De qué forma?	Utilizar los lentes para proyectar las imágenes en movimiento. No todos los niños son visuales, agregar sonido acorde para niños podría ser un valor agregado, pero tener en cuenta a los niños que pueden ser intolerantes a ello.
¿Alguna duda sobre la explicación de la aplicación y su manejo?	Sí se necesita algún tipo de capacitación o si los padres podrían hacer uso de la aplicación.

Elaboración propia

La profesora Ojeda aprueba el uso de una aplicación basada en realidad aumentada para aumentar el vocabulario y mejorar la comprensión lectora. Menciona que durante el primer bimestre es cuando se centran en enseñar palabras con sus significados y se va reforzando durante el resto del año, por lo que este tipo de aplicación sería muy útil durante todo el año escolar.

## 6.5 Pruebas de complejidad

### 6.5.1 Dispositivos

Antes de realizar la validación con la población se tomó en cuenta probar el uso de la herramienta en diferentes dispositivos, como tabletas que varían en tamaño; sin embargo, al momento de la implementación de la solución y por la mayoría de respuestas por parte de los alumnos durante la validación poblacional se decidió que el uso de la aplicación solo podría ser posible con un celular, debido a que los alumnos requirieron de ayuda al momento de hacer uso de la herramienta.

### 6.5.2 Luz

La luz es un factor importante para que la aplicación pueda reconocer los *image target* y así mostrar el resultado en 3D. Usando un luxómetro, instrumento para medir la iluminación real, se ha medido en diferentes niveles de iluminación para obtener el nivel mínimo necesario para que se reconozcan los *image target*.

Tabla 13

*Resultados de prueba de niveles de iluminación*

Nivel de iluminación	15 lx	35 lx	46 lx	65 lx	82 lx	160 lx	356 lx
Reconocimiento	No	Demora	Demora	Sí	Sí	Sí	Sí

Elaboración propia

Los resultados obtenidos muestran que si bien en las pruebas con un nivel de iluminación de 35 lx y 46 lx la aplicación es capaz de reconocer el *image target*, este se demora y el resultado en 3D tiende a desaparecer y aparecer. A partir de 65 lx es un nivel de iluminación ideal.

### 6.5.3 Batería del dispositivo

Se usó la aplicación teniendo al dispositivo con el brillo máximo posible por 6 minutos 18 segundos, tiempo promedio que se tomó el grupo experimental en leer los capítulos indicados. La aplicación estuvo en primer plano en todo momento. Con las variables del brillo y tiempo, el porcentaje de batería reducido fue de 48 %.

## 7. CONCLUSIONES

La investigación comienza con el objetivo de desarrollar una aplicación basada en realidad aumentada para incrementar el vocabulario y con ello mejorar la comprensión lectora. Dicho objetivo se alcanzó de forma satisfactoria, con el uso de objetos en 3D hay un aumento de 0,8 puntos entre el puntaje promedio que hubieran conseguido sin el uso de la *app* y el puntaje promedio conseguido al usar la *app*. Además, hay una diferencia de 0,3 entre el puntaje promedio del grupo de control y el grupo experimental.

Respecto a la investigación, se debe tener en cuenta a la población que va dirigida la solución, pues las herramientas y los dispositivos a utilizar pueden no ser los adecuados. La solución debe adaptarse a la población, no al revés. Para esta investigación se tomó en cuenta que la población de interés son alumnos de segundo grado de primaria, por lo que el tipo de dispositivo utilizado se adaptó a lo que ellos son capaces de utilizar, es por eso que usar un dispositivo más grande como un *ipad* estaba fuera de límites. También los objetos 3D fueron adaptados para que sean adecuados para los niños, que no les produzca reacciones negativas.

## 8. RECOMENDACIONES PARA TRABAJOS FUTUROS

Para el caso del uso de la realidad aumentada, la investigación puede ser extendida utilizando otro tipo de palabras, como verbos o sustantivos que no son capaces de ser explicados con un solo objeto. Además, ciertas palabras podrían requerir un adicional auditivo, como el caso de animales; sin embargo, hay que tener en cuenta los comentarios en la validación por parte de expertos.

Es recomendable escoger como máximo tres palabras por página, debido a que implica la posibilidad que al alumno le llame más la atención los marcadores que el texto y pierda concentración, al tener que estar usando el dispositivo de forma seguida. Además, se pueden sentir confundidos al ver varios ítems desconocidos.

## REFERENCIAS

- Azuma, R. T. (2017) Making augmented reality a reality. *Imaging and Applied Optics 2017 (3D, AIO, COSI, IS, MATH, pCAOP)*. OSA Technical Digest. Optical Society of America, paper JTu1F.1.
- Bacca, J., Baldiris, S., Fabregat, R., Graf, S., y Kinshuk. (2014). Augmented reality trends in education: A systematic review of research and applications. *Journal of Educational Technology & Society*, 17(4), 133-149. Recuperado de <https://www.jstor.org/stable/jeductechsoci.17.4.133>

- Barba Martín, J. J. (2010). Diferencias entre el aprendizaje cooperativo y la asignación de tareas en la escuela rural. Comparación de dos estudios de casos en una unidad didáctica de *acrosport* en segundo ciclo de primaria. *Retos. Nuevas Tendencias en Educación Física, Deporte y Recreación*, 18(2006), 14-18.
- Barba Vera, R., Yasaca Pucuna, S., y Manosalvas Vaca, C. (2015). Impacto de la realidad aumentada móvil en el proceso enseñanza-aprendizaje de estudiantes universitarios del área de medicina. En AIDIPE (Ed.), *Investigar con y para la sociedad* (vol. 3, pp. 1421-1429). Cádiz: Bubok. Recuperado de <http://aidipe2015.aidipe.org>
- Berndtsson, M., Hansson, J., Olsson, B., y Lundell, B. (2008). *Thesis projects. A guide for students in computer science and information systems*. Londres: Springer Science and Business Media.
- Braslavsky, B. (2005). *Enseñar a entender lo que se lee. La alfabetización en la familia y en la escuela*. Buenos Aires: Fondo de Cultura Económica.
- Bravo, L., Villalón, M., y Orellano, E. (2005). El rendimiento en la lectura en el primer ciclo básico y algunos procesos psicolingüísticos de ingreso. *Boletín de Investigación Educativa*, 20(1), 51-63.
- Cabero Almenara, J., Leiva Olivencia, J. J., Moreno Martínez, N. M., Barroso Osuna, J., y López Meneses, E. (2016). Realidad aumentada y educación. *Innovación en contextos formativos*. Barcelona: Octaedro.
- Camilli Trujillo, C., López Gómez, E., y Barceló Cerdá, M. L. (2012). Eficacia del aprendizaje cooperativo en comparación con situaciones competitivas o individuales. Su aplicación en la tecnología: Una revisión sistemática. *Enseñanza y Teaching*, 30(2), 81-103.
- Canet-Juric, L., Burin, D., Andrés, M. L., y Urquijo, S. (2013). Perfil cognitivo de niños con rendimientos bajos en comprensión lectora. *Anales de Psicología*, 29(3), 996-1005.
- Carvajal, G. (2015). Formar lectores en la era digital. *Revista Chilena de Literatura*, (94), 29-300.
- Defior, S. (1996). Una clasificación de las tareas utilizadas en la evaluación de las habilidades fonológicas y algunas ideas para su mejora. *Infancia y Aprendizaje*, 19(73), 49-63.
- Fernández, M. M. (2004). Cómo evaluar la comprensión lectora: alternativas y limitaciones. *Revista de Educación*, (335), 415-127.
- Fundación Telefónica, F. (2011). *Realidad aumentada: una nueva lente para ver el mundo*. Madrid: Editorial Planeta y Fundación Telefónica.
- Gutierrez-Braojos, C., y Salmerón Pérez, H. (2012). Estrategias de comprensión lectora: enseñanza y evaluación en educación primaria. *Profesorado. Revista de Currículum y Formación de Profesorado*, 16(1), 183-202.

- Hedley, N. R., Billinghamurst, M., Postner, L., May, R., y Kato, H. (2002). Explorations in the use of augmented reality for geographic visualization. *Presence: Teleoperators & Virtual Environments*, 11(2), 119-133.
- Hernández Sampieri, C. R., Fernández Collado, C., y Baptista Lucio, M. del P. (2014). *Metodología de la investigación* (6.ª ed.). México D.F.: McGraw-Hill.
- Kaufmann, H., y Schmalstieg, D. (2002). Mathematics and geometry education with collaborative augmented reality. *ACM SIGGRAPH 2002 conference abstracts and applications (SIGGRAPH '02)*. New York: Association for Computing Machinery, 37-41. doi:10.1145/1242073.1242086
- Martín-Sabarís, R. M., Brossy-Scaringi, G. (2017). La realidad aumentada aplicada al aprendizaje en personas con síndrome de Down: un estudio exploratorio. *Revista Latina de Comunicación Social*, 72, 737-750.
- Martínez, R. D., Astiz, M. S., Medina, P. A., Montero, Y. H., y Pedrosa, M. E. (1998). Alternativas para la utilización del hipertexto en el ámbito escolar. *IE Comunicaciones: Revista Iberoamericana de Informática Educativa*, (11).
- Ministerio de Educación del Perú (Minedu). (2015a). ¿Qué logran nuestros estudiantes en la ECE? Recuperado de [http://umc.minedu.gob.pe/wpcontent/uploads/2016/03/Informe\\_IE\\_ECE\\_2015.pdf](http://umc.minedu.gob.pe/wpcontent/uploads/2016/03/Informe_IE_ECE_2015.pdf)
- Ministerio de Educación del Perú (Minedu). (2015b). *Aprendizajes de primero a sexto de primaria en lectura y matemática: un estudio longitudinal en instituciones educativas de Lima Metropolitana*. Lima: Gráfica Técnica S.R.L.
- Ministerio de Educación del Perú (Minedu). (2016a). Currículo Nacional de la Educación Básica. Recuperado de <http://www.minedu.gob.pe/curriculo/pdf/curriculo-nacional-2016-2.pdf>
- Ministerio de Educación del Perú (Minedu). (2016b). Evaluación PISA 2015. Primeros resultados. Recuperado de <http://umc.minedu.gob.pe/wp-content/uploads/2016/12/presentacion-web-PISA.pdf>
- Ministerio de Educación del Perú (Minedu). (2017). ¿Cuánto aprenden nuestros estudiantes? Resultados de la ECE 2016. Recuperado de <http://umc.minedu.gob.pe/wp-content/uploads/2017/04/Resultados-Nacionales2016.pdf>
- Ministerio de Educación del Perú (Minedu). (s. f.). *Demostrando lo que aprendimos: Primer trimestre segundo grado de primaria*. Recuperado de [http://www.perueduca.pe/recursosedu/cuadernillos/primaria/comunicacion/cuadernillo1\\_comunicacion\\_1er\\_trimestre\\_2do\\_grado.pdf](http://www.perueduca.pe/recursosedu/cuadernillos/primaria/comunicacion/cuadernillo1_comunicacion_1er_trimestre_2do_grado.pdf)
- Nava, M. R. Z., González, C. F. M., Galicia, H. A., y Flores, J. M. P. (2016). Marcadores para la realidad aumentada para fines educativos. *ReCIBE*, 2(3).



- OECD (2016). PISA Resultados clave, p. 5. Recuperado de <https://www.oecd.org/pisa/pisa-2015-results-in-focus-ESP.pdf>
- Parra Valcarce, D., Edo Bolós, C., y Marcos Recio, J. C. (2017). Análisis de la aplicación de las tecnologías de realidad aumentada en los procesos productivos de los medios de comunicación españoles. *Revista Latina de Comunicación Social*, (72), 1670-1688.
- Paz A., y Paz C. (2012). *Plan Lector: De la A a la Z Perú*. Lima: Editorial Everest.
- Prendes Espinosa, C. (2015). Realidad aumentada y educación: análisis de experiencias prácticas. *Píxel-Bit. Revista de Medios y Educación*, 46, 187-203.
- Rebollo Catalán, M. A., García Pérez, R., Buzón García, O., y Barragán Sánchez, R. (2012). Las comunidades virtuales como potencial pedagógico para el aprendizaje colaborativo a través de las TIC. *Enseñanza & Teaching*, 30(2), 105-126.
- Reinoso, R. (2016). Realidad aumentada posibilidades y usos educativos. En Baldiris, S. y otros (eds.). *Recursos educativos aumentados. Una oportunidad para la inclusión* (pp. 8-29). Bogotá: Sello Editorial Tecnológico Comfenalco.
- Ruiz Torres, D. (2013). *La realidad aumentada y su aplicación en el patrimonio cultural*. Gijón: TREA.
- Rueda Ortiz, R. (2001). *Evaluación de hipertextos. Perspectivas de diseño e investigación educativa*. Bogotá: Universidad Central Colombia, 275-285.
- Salazar Correa, N. A. (2018). Aplicación del Programa “Mis lecturas favoritas” en el desarrollo del nivel de comprensión lectora en las alumnas del sexto grado de Educación Primaria de la Institución Educativa N.º 11523 de Pucalá - Lambayeque.
- Sampieri, R. H., Collado, C. F., Lucio, P. B., Valencia, S. M., y Torres, C. P. M. (1998). *Metodología de la investigación*. (vol. 6). México, D. F.: McGraw-Hill.
- Schunk, D. H. (2012). *Teoría del aprendizaje: Una perspectiva educativa* (6.ª ed.). México: Pearson.
- Soriano, M., Vidal-Abarca, E., y Miranda, A. (2003). Comparing two instruction procedures for text comprehension and text learning: Direct instruction and reciprocal teaching. *Journal for the Study of Education and Development*, 19(74), 57-65. doi:10.1174/021037096763000781
- Tabares Higueta, L. X. (2017). El hipertexto como herramienta educativa: un recorrido conceptual. *Revista Q*, 3(5).
- Thorne, C., Morla, K., Uccelli, P., Nakano, T., Mauchi, B., Landeo, L., ... y Huerta, R. (2013). Efecto de una plataforma virtual en comprensión de lectura y vocabulario: Una alternativa para mejorar las capacidades lectoras en primaria. *Revista de Psicología*. 31(1), 3-35.

- Unicef (2003). Guía metodológica. Video validación de materiales IEC. Recuperado de <https://www.unicef.org/peru/informes/gu%C3%ADa-metodol%C3%B3gica-video-validaci%C3%B3n-demateriales>
- UMC - Minedu. (2016). Resultados de la evaluación censal de estudiantes. Recuperado de <http://umc.minedu.gob.pe/wp-content/uploads/2017/04/presentacion-ECE-2016.pdf>
- Vallés Arándiga, A. (2005). Comprensión lectora y procesos psicológicos. *Liberabit*, 11(11) 49-61.
- Vallés Arándiga, A., y Vallés Tortosa, C. (2006). *Comprensión lectora y estudio. Intervención psicopedagógica*. Valencia: Promolibro.
- Wittrock, M. C. (1985) *La investigación de la enseñanza, II. Métodos cualitativos y de observación*. Barcelona: Paidós Educador.

# Metrics to Understand Future Maintenance Effort Required of a Complicated Source Code

Michael Dorin

mike.dorin@stthomas.edu / Universität Würzburg, Würzburg, Germany

Sergio Montenegro

sergio.montenegro@uni-wuerzburg.de / Universität Würzburg, Germany

Recepción: 17-6-2019 / Aceptación: 9-7-2019

**ABSTRACT.** An enduring engineering problem is the creation of a source code too complicated for humans to review and understand. A consequence of a complicated source code is that it requires more effort to be implemented and maintained. Exacerbating the problem is a lack of a proper understanding of exactly what the words “complicated” and “complex” mean, as the definitions of these words are often misconstrued. Some systems are indeed inherently complex, but this does not mean they must be complicated. In our research, several open-source projects were evaluated using software metrics to map the complicatedness of a source code with the ongoing effort to sustain the project. The results of our research show that a relationship exists between a complicated source code and the maintenance effort. It is clear that adhering to proper coding practices and avoiding a complicated code can result in a much more manageable future maintenance effort.

**KEYWORDS:** analysis, metrics, complex, complicated, source code

## Métricas para comprender el esfuerzo de mantenimiento futuro requerido de un código fuente complicado

**RESUMEN.** Un problema de ingeniería duradero es la creación de código fuente demasiado complicado para que los humanos lo revisen y comprendan. Una consecuencia del código fuente complicado es que requiere más esfuerzo para implementar y mantener. Lo que exacerbaba el problema es la falta de una comprensión adecuada de lo que significan exactamente las palabras “complicado” y “complejo”, ya que estas definiciones a menudo se interpretan mal. Algunos sistemas son realmente intrínsecamente complejos, pero esto no significa que deban ser complicados. En nuestra investigación, se evaluaron varios proyectos de código abierto utilizando métricas de software para mapear la complejidad del código fuente con el esfuerzo continuo para mantener el proyecto. Los resultados de nuestra investigación muestran que existe una relación entre el código fuente complicado y el esfuerzo de mantenimiento. Está claro que adherirse a las prácticas de codificación adecuadas y evitar el código complicado puede resultar en un esfuerzo de mantenimiento futuro mucho más manejable.

Palabras clave: análisis, métricas, código fuente complejo, complicado

## 1. INTRODUCTION

Life on earth presents humans with some very complex and complicated concepts. It is easy for people to suppose that cancer and quantum mechanics are both complex and complicated, as they are defined like that by nature. The best humans can do is to make simplified model representations of complex physical phenomena. Software development is unique in the world of science and math, such that developers can choose if our work is going to be complicated.

Our research begins by examining and affirming the definitions of *complicated* and *complex*. Afterward, the consequences of *complicated* are measured with a thorough examination of the bug reports and the source code of prominent software projects. The analysis focused on the components of a *LAMP* server (Linux, Apache, MySQL, PHP). The results of this work show that when a source code becomes complicated, it requires more effort to sustain. By establishing the indicators of *complicated* and *complex*, a protocol can be defined to reduce the development of complicated software. As mentioned before, regardless of problem complexity, developers can choose whether their code will be complicated.

## 2. BACKGROUND

### *Complicated and Complex*

The words *complicated* and *complex* are often misconstrued, and it is possible to have an understanding that the two words mean the same thing. This fact becomes evident when we review standard definitions of the words, as it can be challenging to notice the differences. For example, a standard definition of *complicated* indicates that something complicated is composed of interconnected parts and is difficult to analyze (Mish, 2004). Similarly, something complex is “composed of many interconnected parts; compound; composite” (Mish, 2004). These definitions almost sound identical.

The etymology of the word *complicated* shows the Latin root “plic,” which means “to fold”. The word *complex* has the Latin root “plex,” which means “to weave”. Something that is woven has many strands, but the strands can be well organized. When “fold” is used as a suffix, it means a “multiple”. Something complicated can have fragments folded in, making it conceptually tricky to understand (Lissack, 1999).

Finally, when we consider whether something is complicated, we must acknowledge that, many times, something is complicated when it is not understood. Upon understanding a topic, it is often no longer complicated. With this in mind, consider that the number of people involved also plays a role in creating complicatedness. Having more people involved means that it is more likely that a single person will not understand it.

### *Git*

The use of features found in Git was an essential part of this research. Git is a prevalent and heavily utilized version control system developed in 2005 by Linus Torvalds (Spinellis, 2012). Logs created by Git were used extensively and were examined using the Git log command. A python script was used to parse the logs which focused on Git commits. A Git commit means that a file was created or updated and stored in a local Git repository (Torek, 2019).

### *Static Analysis*

Static analysis examines a program source code without actually executing the program. One of the first tools for performing static analysis was *Lint*, created in 1978 (Bardas, 2010). Static analysis has improved considerably since then and bugs flagged by the original Lint are now included in the compiler output. Many static analysis tools exist at this time, and the various tools specialize in different areas (Bardas, 2010). This study used the Cppcheck tool (Marjamäki, 2019) as it looks for multiple types of errors and is easy to run immediately without integrating it into a functioning build.

### *Measurements and Metrics*

Attempting to understand complications and *complexity* is not a new concept, and several metrics have been invented throughout the decades to quantify these subjects. The first metric would merely be counting lines of code and reporting using units such as LOC (lines of code) and KLOC (thousands of lines of code). In 1976, Thomas McCabe developed the cyclomatic complexity metric, which is a measure of the number of linearly independent paths through a program (McCabe 1976). In 1977, Maurice Halstead published complexity metrics based on distinct operators, distinct operands, the total number of operators, and the total number of operands (Halstead, 1977). Using Halstead attributes, a value for difficulty and effort is made. Both Halstead and McCabe metrics have been both praised and criticized, but they are relatively simple and practical contributions to the software engineering community. More advanced metrics exist, but as this research is focused on identifying complexity and complicatedness, we have decided to use uncomplex and uncomplicated metrics as part of our work.

## **3. RELATED WORK**

### *Existing Literature*

Several research studies that consider the human aspects of a complicated source code have been conducted. Daniel Sturtevant has done an excellent work in his thesis for MIT describing

the impact of software architectural design, and connecting it to employee productivity and staff turnover. He worked with a professional organization, and included a case study of a development project which shows that the structure of a software system does impact productivity and employee retention. Sturtevant points out that further research on different types of projects is warranted, and that there are many opportunities for this (Sturtevant, 2013).

Human considerations are further explored in the paper from Foucault et al. (2015) researching the impact that developer turnover has on software quality. This paper covers open-source software, and demonstrates a link between developer turnover and the number of bugs found in the project (Foucault, 2015). Interestingly, Foucault and Sturtevant do not entirely agree on the aspect of whether poorly designed software drives developers away or whether the turnover creates a poorly designed software.

The turnover of developers might also be a reason for code churn as described in the paper *Use of Relative Code Churn Measurements to Predict System Defect Density* (Nagappan, 2005). Code churn is a metric based on how much the code is changed, and this metric has been shown to be an indicator of faults. It is suggested that more recent code is faultier than the original code (Nagappan, 2015).

Other papers analyzed measurable aspects of software itself, rather than human influence. In the article *Reexamining the Fault Density - Component Size Connection*, Les Hatton discusses the relationship with the size of the program and the number of faults. Lines of code (LOC) has always been a good predictor of fault opportunities, and it is intuitively apparent why (Hatton, 1997). A four-line “Hello World” program will have fewer opportunities for faults than a four thousand-line program. Having more lines offers more opportunities for mistakes. Though LOC is a good indicator of bug opportunities, it also does not help you narrow down specific locations where faults may be present.

### *Previous Project Work*

In a previous work on this project, we determined which programming practices software was the most difficult for engineers. We discovered what engineers misinterpret and what software engineers find displeasing to review. Also, a consistent unpleasant to review code had a higher Halstead and cyclomatic complexity. Stylistic constructs that programmers generally have difficulty evaluating are shown in Table 1 (Dorin, 2018).

Table 1  
*Unpleasant Styles (Dorin, 2018)*

Style Name
Avoid too deep blocks
Do not write over 120 columns per line
Matching braces should be in the same column
Use less than five parameters
Use braces even for one statement
Indent blocks inside a function

For measuring project conformance with stylistic rules, the `nsiqcppstyle` tool (Yoon, 2014) was used. Using this previously developed knowledge in a complex and complicated code, we are now expanding the work to develop indicators to relate a complicated source code to this ongoing effort.

#### 4. METHODOLOGY

Existing literature covers a lot of important subjects, but many papers seem to be specific to particular cases and specific environments. In this project, we reviewed several active open-source projects maintained by Git. Popular projects were preferred as they gave the most opportunities for users to report bugs. GitHub.com was used as a source for all the projects. The four cornerstone projects selected were Linux, Apache, MySQL, and PHP. The remaining projects were found based on a report of active projects on GitHub. Once projects were found, the following steps were performed:

1. *Calculation of thousands of lines of code (KLOC) for C, C++, and headers.* The amount of information given to a human to process is a factor in how much can be understood. For each project, for each source related file, we used the Lizard tool (Yin, 2019) to calculate the lines of code in both source files and headers.
2. *Calculation of Halstead for C, C++, and headers.* The Halstead metric, which evaluates complexity based on distinct operators, distinct operands, the total number of operators, and the total number of operands, was also used in this study. For each project, we used a program called commented code detector (Borowiec, 2014) to measure the Halstead metrics.

3. *Calculation of McCabe cyclomatic complexity for C, C++, and headers.* The McCabe cyclomatic complexity metric is based on the number of linearly independent paths through code. The cyclomatic complexity was used because more paths require more effort to understand. The Lizard tool (Yin, 2019) was also used to gather information on cyclomatic complexity for each project.
4. *Determination of “bug”-sized changes.* Several small utilities were written to analyze the logs maintained by Git. When searching for keywords, we determined the typical number of lines of code found in bugs and thus classified bugs by size. We classified tiny bugs as “gnats,” as they are important for consideration since numerous small bugs take resources away from developing essential features. The gnat-sized bugs constitute 50% of the Git check-ins, which is a good indication of how bothersome they are.
5. *Measurement of conformance to stylistic rules.* From our previous work, we determined what was essential and what rules had the most impact on programmer understanding. As with our past projects, we used the `nsiqcppstyle` tool to measure stylistic rule conformance.
6. *Measurement of basic coupling between modules.* Though many utilities exist for measuring coupling, they were not practical for our use. For coupling, we created a straightforward metric which we called *Sheficom*. The *Sheficom* metric computes how many external modules are coupled using a count of the number of headers in the module. *Sheficom* is based on the assumption that, if a module is not coupled, it would not need to include a header.
7. *Measurement of the percentage of bug-related changes.* The goal of this effort was to measure how much work in a project was dedicated to bug fixes. By using the technique outlined in Item 4 to identify bugs, the percentage of bug-related changes was measured. We counted the aggregate number of complete bugs as well as the total number of lines changed for fixing bugs.
8. *Count of the number of authors.* As mentioned, one of the natures of being complicated is the absence of understanding. More programmers mean more people are required to understand the source code. Depending on the communication paths and capabilities of the authors, understanding could be impacted.
9. *Performance of static analysis of modules.* The `Cppcheck` tool (Marjamäki, 2019) was used to perform static analysis on the source code. This tool was chosen as it does not require performing any special actions or creating special configurations to analyze the source code, but it will provide an estimate of errors which exist in the project.



Table 2

Metric	Software Complexity	Human Complication
Size		X
Style		X
Sheficom (coupling)	X	
Author Turnover		X
Halstead	X	
Cyclomatic	X	

These metrics were used to get a picture of the effort needed for ongoing maintenance work. All of the metrics, in one way or another, contribute to impacting effort. However, some metrics have a more significant influence on what humans perceive as complicated (see Table 1). For example, in the new Sheficom metric, tightly coupled modules take a human longer to review and therefore more time to understand. It is plausible that a human, with enough time, could understand what is going on. At that point, though it remains complex, the code is no longer complicated. Author turnover was also essential to measure, as new authors require time to build up an understanding of the existing code.

## 5. RESULTS

### *Linux*

Linux is a very well-established operating system and, for this study, we performed measurements every year for ten years, starting in 2008. See Table 3 and Figure 2 for a summary of the results. When examining the coding style, it seems to conform less year by year (see Figure 2). Improvements in module coupling and cyclomatic complexity are apparent. However, the size of the project grows linearly over time, and the number of static analysis detected bugs per snapshot also increases over time. More than forty percent of the effort to continue Linux seems to be changes related to bugs, as shown in Table 3. One more thing to bear in mind is that Linux has a lot of authors working on the project, which will contribute to the project being more complicated. It seems that a lot of authors and a lot of code lead to misunderstandings, and a lot of effort is spent fixing bugs rather than adding new functionality.

Metric	Total Change	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Style		43 %	43 %	43 %	42 %	42 %	41 %	40 %	40 %	40 %	40 %	39 %
Sheficom		12.55	12.36	12.6	12.78	12.58	12.42	12.26	12.16	12.1	12.15	11.94
Halstead		51.46	51.06	50.48	51.3	51.54	52.12	54.24	55.19	55.72	55.2	57.96
Cyclomatic		4.5	4.5	4.4	4.4	4.4	4.3	4.3	4.3	4.3	4.3	4.3
KLOC		5646	7045	79.66	8673	9521	10459	11090	11941	12938	14094	14519
Existing Authors		48 %	51 %	54 %	55 %	58 %	60 %	57 %	59 %	62 %	61 %	64 %
Bug Changes		51 %	41 %	41 %	43 %	46 %	52 %	47 %	40 %	46 %	55 %	69 %
Bug Commits		45 %	43 %	42 %	39 %	39 %	39 %	40 %	41 %	43 %	43 %	43 %
Cppcheck		6247	6219	10371	12889	14406	15551	17127	16530	18317	19979	21494

Figure 1. Linux Year-by-Year Results

### Apache

The results of the analysis of Apache are listed in Figure 3 and Table 4. As far as significant projects go, Apache does an outstanding job with the level of effort dedicated to bug fixes versus improvements (~32% for bug fixes). A positive aspect of Apache is that many authors remain from release to release (~85%), which means understanding can remain high, thus reducing the impact of a complicated code. We can also see improvements to cyclomatic complexity. Though Halstead measurements remain pretty steady, the coupling is pretty constant, and the style is pretty steady. Even the Cppcheck static analysis snapshot remains steady.

Metric	Total Change	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Style		33.3 %	31.6 %	31.0 %	32.3 %	32.3 %	32.7 %	31.9 %	32.6 %	32.1 %	31.8 %	31.0 %
Sheficom		10.7	10.65	10.74	10.99	10.92	11	11.27	11.56	11.88	11.81	12
Halstead		61.47	62.42	63.06	62.94	62.63	63.4	63.67	61.14	62.28	62.76	62.2
Cyclomatic		7.57	6.74	6.75	6.68	6.68	6.63	6.64	6.5	6.44	6.37	6.23
KLOC		143	125	131	139	151	162	170	178	189	196	209
Existing Authors		81 %	85 %	85 %	91 %	75 %	92 %	78 %	77 %	86 %	94 %	93 %
Bug Commits		14 %	43 %	37 %	41 %	98 %	28 %	9 %	18 %	21 %	26 %	26 %
Bug Changes		25 %	30 %	37 %	31 %	35 %	32 %	34 %	30 %	27 %	32 %	31 %
Cppcheck		60	68	58	61	58	64	65	64	64	67	68

Figure 2. Apache Year-by-Year Results

## MySQL

The results for MySQL are listed in Figure 4 and Table 5. MySQL has dramatically increased in the number of lines of code. Note the codebase growth between 2015 and 2018, where a large amount of code was added. Along with the dramatic increase in the number of lines of code, the coupling (Sheficom) has also increased. Also, the problems predicted by static analysis (Cppcheck) have increased, and most importantly, the amount of effort spent on fixing bugs has also increased over time.

Metric	Total Change	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Style		39 %	38 %	38 %	39 %	40 %	41 %	41 %	44 %	45 %	48 %	42 %
Sheficom		6.02	6.07	6.79	7.04	7.76	8.06	8.18	8.66	9.1	13.07	13.67
Halstead		71.08	71.4	71.3	70.6	72.12	69.78	69.13	66.12	64.21	62.15	62.48
Cyclomatic		4.54	4.55	4.54	4.38	4.11	4.09	4.03	3.9	3.6	3.52	3.68
KLOC		904	923	1160	1204	1493	1571	1674	1833	2066	2581	2715
Existing Authors		46 %	58 %	66 %	71 %	66 %	73 %	73 %	67 %	82 %	77 %	92 %
Bug Commits		33 %	37 %	34 %	30 %	35 %	36 %	40 %	37 %	41 %	42 %	42 %
Bug Changes		60 %	50 %	41 %	66 %	59 %	66 %	61 %	65 %	65 %	70 %	95 %
Cppcheck		116	119	124	130	143	224	259	278	312	422	518

Figure 3. MySQL Year-by-Year Results

## PHP

PHP has conditions showing a more complicated code in several areas, as shown in Figure 5. Style conformance has dropped while the number of lines of code has grown. The number of veteran authors has fluctuated, contributing to only a few people having an ongoing understanding of the code. Halstead and cyclomatic complexity both indicate a complicated code in this product, which will also negatively impact new authors joining the project. PHP is spending about half of their time fixing issues rather than adding new features.

Metric	Total Change	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Style		42 %	43 %	44 %	43 %	43 %	43 %	43 %	43 %	43 %	42 %	41 %
Sheficom		6.47	6.95	6.86	7.25	7.39	7.34	7.31	12.16	7.22	7.18	7.18
Halstead		115.1	134.1	131.6	128.2	131.8	127.5	131.6	55.19	134.0	151.9	151.1
Cyclomatic		7.68	7.65	7.69	7.65	7.74	7.82	7.77	4.3	8.27	8.23	8.19
KLOC		600	789	733	745	812	836	900	11941	1036	1183	1234
Existing Authors		59 %	83 %	80 %	84 %	49 %	45 %	53 %	59 %	42 %	45 %	64 %
Bug Changes		41 %	38 %	31 %	64 %	75 %	65 %	77 %	40 %	50 %	79 %	40 %
Bug Commits		54 %	61 %	59 %	59 %	41 %	36 %	37 %	41 %	30 %	25 %	24 %
Cppcheck		466	562	558	640	663	694	692	16530	257	362	370

Figure 4. PHP Year-by-Year Results

*ImageMagic*

Finally, the results for the ImageMagic project were also analyzed and are shown in Figure 6. In this project, style conformance is not that much different from other projects. Measurements like Halstead and cyclomatic complexity are not superb either. However, note the strong connection between existing authors and bug changes. After 2014, a decrease in existing authors coincides with a dramatic increase in bug efforts. The amount of energy spent on bugs impacted the amount of effort on new features as new authors arrived.

Metric	Total Change	2010	2011	2012	2013	2014	2015	2016	2017	2018
Style		40 %	41 %	41 %	40 %	40 %	40 %	41 %	40 %	40 %
Sheficom		17.24	18.13	17.93	18.77	19.05	19.2	19.96	20.04	19.5
Halstead		12.79	12.08	11.47	11.69	11.51	11.43	11.73	11.71	11.3
Cyclomatic		9.1	9.4	9.4	9.4	9.1	9.2	9.3	9.4	9.5
KLOC		287	297	314	315	335	341	342	346	353
Existing Authors		75 %	100 %	100 %	80%	100 %	44 %	14 %	12 %	11 %
Bug Changes		40 %	0 %	1 %	2%	18 %	11 %	8 %	21 %	38 %
Bug Commits		5 %	3 %	27 %	5%	13 %	17 %	27 %	34 %	47 %
Cppcheck		19	24	35	31	26	24	11	11	11

Figure 5. Image Magic Year-by-Year Results

**6. DISCUSSION**

This study used several metrics as a means of identifying how complicated a project’s source code is. The metrics covered different areas related to a complicated code, including author turnover and module coupling. The most significant metrics seem to be the size of the project and how many human contributors there are. All software engineers have had the experience that something is initially complicated, but as time goes by and more effort is applied, the project gets less complicated. The metrics related to software complexity, as shown in Table 1, seemed to be consistent year over year on most of the projects. More successful projects maintain veteran programmers on their teams, as it would seem that new authors are not as effective. Complexity metrics, as shown in Table 1, are a harbinger of how long it may take new authors to understand the code and be productive. Projects with veteran programmers consistently show that less time and effort is required for addressing bugs in the maintenance process.

Table 3  
*Average Results*

Category	Linux	Apache	MySQL	PHP	ImageMagic
Lines of Code	14,510,383	208,994	2,714,775	1,233,536	353,088
Lines of Code (average by year)	10,354,856	162,894	1,647,628	892,846	325,549
Style Conformance	41%	32%	42.84%	42.57%	40.33%
Cyclomatic Complexity	4.36	6.7	4.1	7.9	9.31
Halstead Difficulty	53.69	62.55	68.20	133.84	11.74
Halstead Effort	19,878.92	24,169.54	39,801	65,720	47,475.5
Coupling (Sheficom)	12.35	11.23	8.58	7.13	18.66
Authors Year-by-Year					
Average Number	3,336	23	127	88	9
Average Veterans	57%	85%	70%	41.82%	60%
Bug Effort (commits vs. changes)					
Commit Count	41% bugs	31.22% bugs	37% bugs	47% bugs	19.63% bugs
Changes	48% bugs	32.4% bugs	63% bugs	56% bugs	15.43% bugs
Gnat-Sized Bug	15 lines	12 lines	25 lines	9 lines	7 lines
Static Analysis Error Snapshot	14,675	64	240	500	21

## 7. RISKS TO VALIDITY

Several difficulties had to be addressed when doing this project. One difficulty was ascertaining the number of bugs in a project. Often developers find and fix bugs without mentioning them in bug tracking systems. It is also possible that a single “Git commit” has multiple bug fixes included. To overcome this, we evaluated bug-sized fixes and compared them to “Git commits” in general. We may be overestimating or underestimating the bug count. Another area of risk is based on the tools selected. These tools were all available open source and easy to find. However, an exhaustive test on the tools was not performed as part of this study.

Finally, an understanding of the terms *complex* and *complicated* is difficult to arrive at, as humans will all perceive these concepts differently. Something may be thought complicated, but after some time and understanding, the judgment of complicated no longer seems accurate.

## 8. CONCLUSION

The paper shows that there are consequences to having a complicated code. Many aspects play a part in how complicated a source code is, but the most significant contributors seem to be the size of the project and the number of authors participating. More lines of code mean more to understand. The more authors who participate in a project, the more people who need to understand its code. These factors contribute to an increased amount of time spent fixing bugs as opposed to improving the product.

Simple metrics, as used in this paper, can be employed to identify complicated areas of a project. Areas that are identified as complicated by the metrics imply it will take longer to understand those portions of the code. The longer it takes to understand, the more effort required for bug fixing. If a code is hard to understand, new authors will not be able to come up to speed quickly, and existing authors may not wish to remain on the project. Projects need to maintain loyal authors or spend more time on fixing problems instead of adding features.

Areas for future study include more effort mapping static analysis back to a complicated code and measuring how many identified areas remain in a delivered product. The second area of prospective study would be to focus on projects which follow all the rules established and verify if such projects truly have fewer errors than those which ignore these rules.

## REFERENCES

- Bardas, A. G. (2010). Static code analysis. *Journal of Information Systems & Operations Management*, 4(2), 99-107.
- Borowiec, D. (2014). Commented Code Detector. Retrieved June 16, 2019, from <https://github.com/dborowiec/commentedCodeDetector>
- Dorin, M. (2018, May). Coding for inspections and reviews. In *Proceedings of the 19th International Conference on Agile Software Development: Companion* (p. 34). ACM.
- Foucault, M., Palyart, M., Blanc, X., Murphy, G. C., & Falleri, J. R. (2015, August). Impact of developer turnover on quality in open-source software. In *Proceedings of the 2015 10th Joint Meeting on Foundations of Software Engineering* (pp. 829-841). ACM.

- GitHub. (2019). Build software better, together. Retrieved June 16, from <https://github.com/>
- Halstead, M. H. (1977). *Elements of software science* (Vol. 7, p. 127). New York: Elsevier.
- Hatton, L. (1997). Reexamining the fault density component size connection. *IEEE software*, 14(2), 89-97.
- Kobashi, Y., Fukuda, M., Yoshida, K., Miyashita, N., Niki, Y., & Oka, M. (2006). Chronic necrotizing pulmonary aspergillosis as a complication of pulmonary Mycobacterium avium complex disease. *Respirology*, 11(6), 809-813.
- Lissack, M., & Roos, J. (1999). *The next common sense: Mastering corporate complexity through coherence*. Nicholas Brealey Publishing.
- Manser, M. (2004). *Good Word Guide*. Bloomsbury Publishing UK.
- Marjamäki, D. (2019). Cppcheck. Retrieved June 16, 2019, from <http://cppcheck.sourceforge.net>
- McCabe, T. J. (1976). A complexity measure. *IEEE Transactions on software Engineering*, (4), 308-320.
- Mish, F. C. (Ed.). (2004). *Merriam-Webster's collegiate dictionary*. Merriam-Webster.
- Nagappan, N., & Ball, T. (2005, May). Use of relative code churn measures to predict system defect density. In *Proceedings of the 27th international conference on Software engineering* (pp. 284-292). ACM.
- Spinellis, D. (2012). Git. *IEEE software*, 29(3), 100-101.
- Sturtevant, D. J. (2013). *System design and the cost of architectural complexity* (Doctoral dissertation, Massachusetts Institute of Technology).
- Torek, A. C. Distributed Version Control With Git And Mercurial. Retrieved June 16, 2019, from <http://web.torek.net/torek/tmp/book.pdf>
- Yin, T. (2019). Lizard. Retrieved June 16, 2019, from <https://github.com/terryyin/lizard>
- Yoon and K. Tyagi. (2014) nsiqcppstyle. Retrieved June 16, 2019, from <https://github.com/kunaltyagi/nsiqcppstyle>





# Red neuronal convolucional para la detección de aves exóticas en peligro de extinción

Guillermo Eduardo Narváez  
guillermo.narvaez.21.95@gmail.com / Universidad de Lima, Perú

Recepción: 17-6-2019 / Aceptación: 8-8-2019

**RESUMEN.** En los últimos años, los índices de comercialización ilegal de animales se han incrementado debido a la falta de conciencia sobre el cuidado y la preservación de nuestro ecosistema. Entre los años 2000 y 2015, 67 749 animales silvestres fueron extraídos del Perú de manera ilegal. Entre ellos 29 591 eran aves exóticas (43 %), 26 951 anfibios (40 %), 8600 reptiles (13 %) y 2607 mamíferos (4 %); valorizados en mil dólares, aproximadamente. La propuesta para reducir significativamente estos actos ilícitos se basa en contar con un aplicativo que pueda identificar a las especies que se intentan comercializar ilegalmente. Dicho aplicativo hará uso de las arquitecturas de las redes neuronales convolucionales denominada VGGnet16, la cual permitirá identificar correctamente al animal. Se ha trabajado con una base de datos de animales reales obteniéndose una efectividad del 89 %.

**PALABRAS CLAVE:** reconocimiento de objetos, redes neuronales convolucionales, comercio de vida silvestre

## Convolutional Neural Network for Detecting Endangered Exotic Birds

**ABSTRACT.** In the last years, the rates of illegal wildlife trade have increased due to the lack of awareness for caring and preserving our ecosystem. Between the years 2000 and 2015, sixty-seven thousand seven hundred forty-nine (67,749) wild animals were illegally taken from Peru, including 29,591 exotic birds (43%), 26,951 amphibians (40%), 8,600 reptiles (13%) and 2,607 mammals (4%) valued at approximately USD 1,000. The proposal to significantly reduce these illegal acts is based on having an application that can identify the species attempted to be illegally traded. Such application will use the architectures of convolutional neural networks called VGGNet16, which will allow the correct identification of the animal. The present research used a database of real animals, obtaining an effectiveness of 89%.

**KEYWORDS:** object recognition, convolutional neural networks, wildlife trade

## 1. INTRODUCCIÓN

El tráfico de animales es el tercer comercio ilegal después de las drogas y las armas (Quevans, Falcón y Elías, 2014). Estos actos conllevan a la depredación de la fauna o biodiversidad que para la mayoría de las personas suele pasar desapercibido. Esto tiene como consecuencia la muerte de cientos e incluso miles de animales durante su captura y transporte (Carrión, 2017). Dicha actividad está en crecimiento, no solo a nivel local sino mundial (Quevans *et al.*, 2014).

En la tabla 1 se detalla la finalidad y los países destino donde los animales son comercializados de manera ilegal (Quevans *et al.*, 2014).

Tabla 1  
*Finalidad del tráfico de fauna silvestre en el Perú*

Finalidad	Descripción
Coleccionistas particulares y zoológicos	Europa: Holanda, Bélgica, Austria, Suiza, Francia, Alemania, Italia Reino Unido y España. Asia: Singapur, Hong Kong, Japón y Filipinas. Norteamérica: Estados Unidos y Canadá.
Científicos y la industria biomédica	Los investigadores ilegales están en búsqueda de nuevas especies y sus derivados para la producción de medicamentos.
Comercialización internacional en tiendas de mascotas	Esto incentiva el tráfico ilegal de animales silvestres ya que los precios varían y dependiendo de la especie existe una mayor demanda.
Otras finalidades	Trofeos de caza deportiva, actividades religiosas, etc.

Fuente: Estrategia nacional para reducir el tráfico ilegal de fauna silvestre en el Perú 2017-2027 y su plan de acción 2017-2022. Serfor (2017).

En el Perú existen altos índices de tráfico de animales debido a que en Europa muchas de nuestras especies se encuentran altamente valorizadas (Barletta, Pereira, Robert y Yoguel, 2013). Algunas de dichas especies son el cóndor andino, el gallito de las rocas, loro cabeza roja, entre otros, 8479 aves fueron decomisadas en pésimas condiciones en el período de 2002-2007 en Lima, Perú (Quevans, *et al.*, 2014). Según Carrión (2017), durante el proceso de intervención en las acciones contra el tráfico ilegal y depredación de fauna silvestre se incautaron cacatúas en botellas de plástico (véase la figura 1).



*Figura 1.* Cacatúa dentro de una botella de plástico

Fuente: Carrión (2017)

Lamentablemente, los lugares de control a nivel nacional son deficientes y no existen suficientes especialistas ni agentes ambientales de la policía ecológica; por lo que el proceso de intervención no es efectivo (Carrión, 2017) y no existe ninguna base de datos digital oficial adecuadamente documentada (López Tarabochia, 2016).

Por lo anteriormente expuesto, surge la necesidad de apoyar el proceso de intervención en las acciones contra el tráfico ilegal y depredación de fauna silvestre mediante el desarrollo de un aplicativo que permita identificar correctamente el tipo de ave que se está traficando. Esta identificación está enlazada con un registro de base de datos que brinda la información si es una especie o no en peligro de extinción.

## 2. ESTADO DEL ARTE

### 2.1 Redes neuronales convolucionales para la detección de objetos a través de imágenes

El reconocimiento de objetos consiste en la extracción de parámetros que permiten una identificación e interpretación inconfundible del objeto. El proceso de reconocimiento de objetos

comienza con el preprocesamiento de las imágenes, seguido de un modelo de clasificación previamente entrenado (Rossius, 2013).

Las redes neuronales convolucionales (CNN, por sus siglas en inglés) se han establecido como una poderosa clase de métodos para problemas de reconocimiento de imágenes (Karpathy, Toderici, Shetty, Leung, Sukthankar y Fei-Fei, 2014). Las arquitecturas de las CNN pueden contener hasta sesenta millones de parámetros (miles de imágenes). Esta idea es clave debido a que las capas internas de las CNN pueden actuar como un extractor genérico de representación de imágenes que puede ser entrenado con un conjunto de datos (Oquab, Bottou, Laptev y Sivic, 2014).

### 2.1.1 Aplicación de CNN en animales

La taxonomía es un elemento muy importante para la gestión de la biodiversidad ya que estas informan sobre las características esenciales y los nombres de las especies. Asimismo, permite mejorar los estudios científicos y programas de vigilancia para cuidar el ecosistema.

Una de las aplicaciones fue la identificación de los géneros de las hormigas haciendo uso de las CNN, para lo cual se les utilizó para identificar y clasificar automáticamente a las hormigas con el fin de conocer sobre su taxonomía. Se utilizó un repositorio (<https://antweb.org/>) donde se obtuvieron imágenes de diferentes ángulos de las hormigas (cabeza, perfil dorsal y superior) (véase la figura 2) (Marques *et al.*, 2018).



Figura 2. Ejemplo de las perspectivas de la cabeza, perfil forsal y superior de una especie llamada *Odontomachus chelifer*

Fuente: Marques *et al.* (2018)

La arquitectura utilizada fue AlexNet con 8 capas convolucionales, 5 capas RELU, 3 Max. Pooling y 3 Full Connected (véase la figura 3).

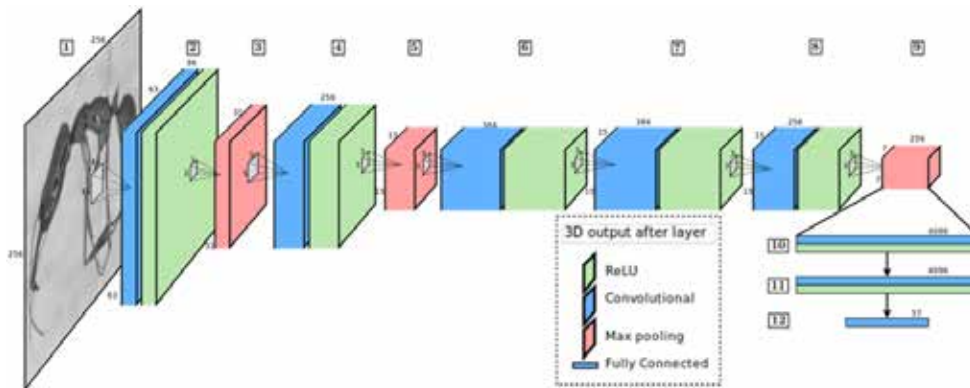


Figura 3. Arquitectura CNN utilizada para realizar la clasificación de imágenes  
Fuente: Marques *et al.* (2018)

Se consideraron diversos clasificadores como exactitud, precisión media y precisión mínima. Los mejores resultados se obtuvieron cuando se trabajaron con imágenes desde el punto de vista de la cabeza. Esto se debe a que la clasificación de vistas de cabeza puede ser más precisa, porque la posición de las piernas de la muestra de perfil y vistas dorsales podría añadir cierto ruido. Por otro lado, los géneros de las hormigas tienen diferencias morfológicas más pronunciado en la cabeza (Marques *et al.*, 2018).

Asimismo, varios estudios demuestran que los virus son los responsables de más de cien enfermedades que componen una enorme carga a nivel mundial (Musso, Rodríguez-Morales, Levi, Cao-Lormeau y Gubler, 2018). Alguno de ellos es el dengue, el chikunguya y zika. Motta y otros (2019) proponen una aplicación basada en CNN para la clasificación de los mosquitos en el campo, llevando a cabo una clasificación morfológica automática. Trabajaron con un conjunto de datos que incluía 4056 imágenes de los mosquitos. Dichas imágenes sirvieron para realizar un entrenamiento en tres arquitecturas de CNN: LeNet, AlexNet y GoogleNet (Motta *et al.*, 2019) (véase la figura 4).

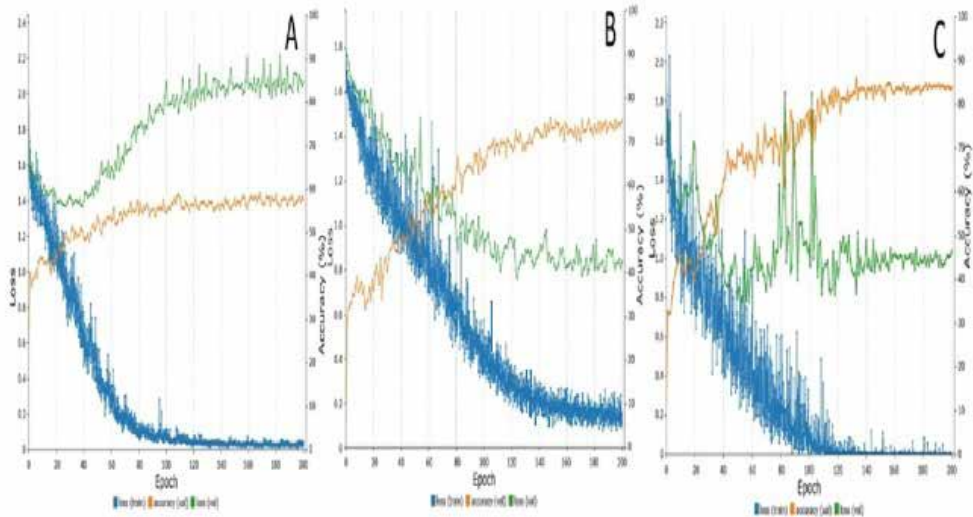


Figura 4. Representación del proceso de capacitación de cada CNN: (1A) LeNet, (1B) AlexNet y (1C) GoogLeNet (línea naranja: validación de precisión; línea verde: validación de pérdida; línea azul: entrenamiento de pérdida). (1A) LeNet (después de 200 épocas): porcentaje de precisión (validación) 57,50 %, pérdida (validación) 2,07, pérdida (entrenamiento) 0,03; (1B) AlexNet (después de 200 épocas): porcentaje de precisión (validación) 74,69 %, pérdida (validación) 0,83, pérdida (entrenamiento) 0,11; (1C) GoogLeNet (después de 200 épocas): porcentaje de precisión (validación) 83,88 %, pérdida (validación) 1,03, pérdida (entrenamiento) 0,00

Fuente: Motta *et al.* (2019)

## 2.2 Aplicaciones móviles con CNN

Durante la vigesimotercera conferencia de reconocimiento de patrones se presentaron estudios sobre el reconocimiento de objetos haciendo uso de celulares centrados en tres arquitecturas: AlexNet, Network In Network y GooLeNet. La aplicación del Museo del Mar en España es un ejemplo de estos propuestos (Tobías, Ducournau, Rousseau, Mercier y Fablet, 2017)

Tobías *et al.* (2017) propuso diferentes cámaras que fueron utilizadas para poder adquirir imágenes con diferentes calidades. Todas las imágenes estuvieron sujetas a un paso de procesamiento previo que incluyó la normalización del tamaño (véase la figura 5).

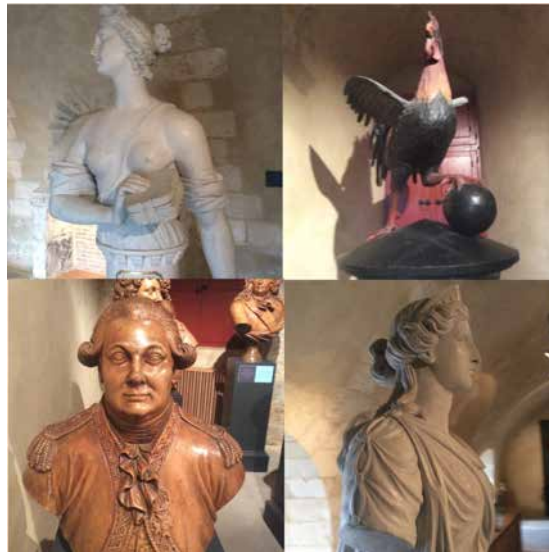


Figura 5. Ejemplo de imágenes de la base de datos del Museo del Mar.  
Fuente: Tobías *et al.* (2017)

Una de las principales ventajas de la utilización de CNN es que los modelos pueden ser desplegados en casi cualquier dispositivo y el tiempo de entrenamiento es directamente proporcional a la cantidad de datos a procesar y el número de iteraciones aplicadas (Tobías *et al.*, 2017). Por ejemplo, en este experimento que se realizó se ejecutó, aproximadamente, en dos a tres días utilizando varios dispositivos: CPU i7.6820HQ, CPU Intel AT 2.70GHz, GPU Nvidia Tesla K80 y un Ipad con un procesador de 1,3 Ghz de doble núcleo (Tobías *et al.*, 2017). Estos datos se aprecian en la tabla 2.

Tabla 2

*Tiempo computacional promedio y memoria de requerimientos profundos*

	GPU	CPU	Mobile	Memory
AlexNet	22,6 ms	339,0 ms	382 ms	230 MB
GoogLeNet	17,6 ms	452,1 ms	992 ms	50 MB

Fuente: Tobías *et al.* (2017)

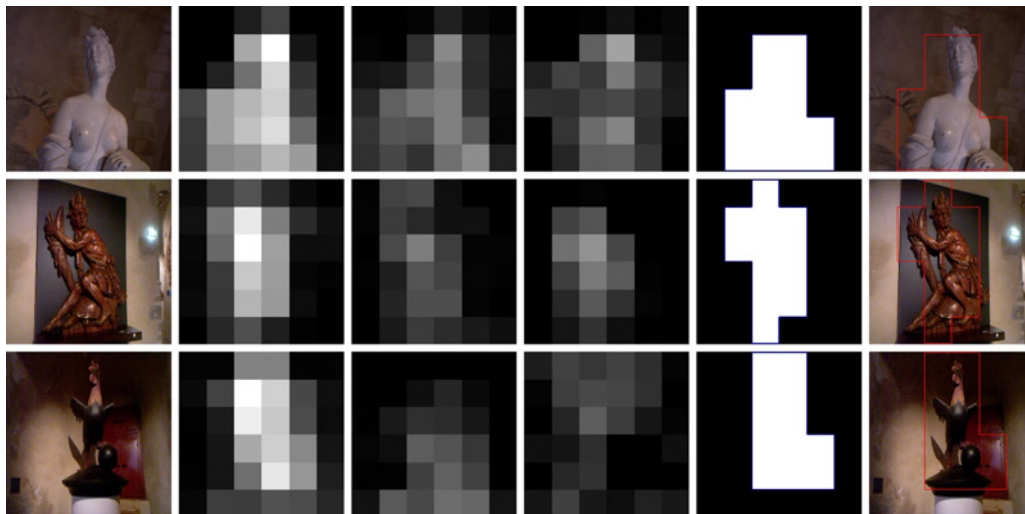


Figura 6. Localización e identificación de objetos, caso de estudio del Museo del Mar  
Fuente: Tobías *et al.* (2017)

### 3. METODOLOGÍA

La metodología de la investigación consiste en que el usuario al abrir el aplicativo activará la cámara del celular de manera automática. Las imágenes capturadas, automáticamente, pasarán por un proceso de identificación de la CNN y si, en el caso, dicha ave está en peligro de extinción se procederá en generar el acta (véase la figura 7).

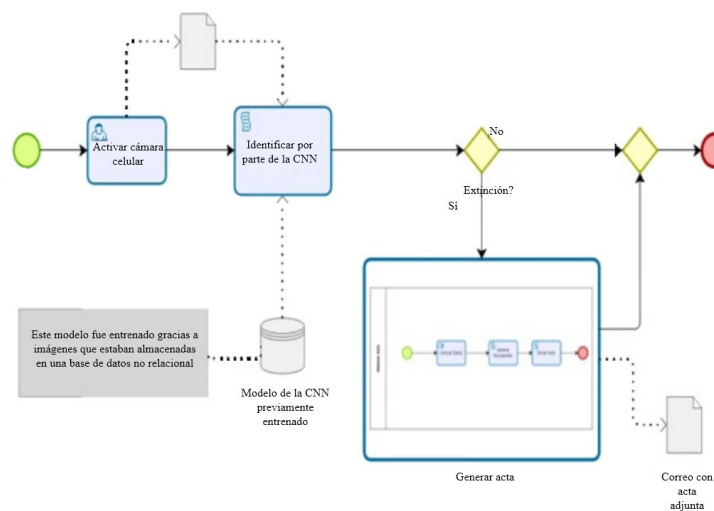


Figura 7. Esquema de la propuesta de solución  
Elaboración propia



### 3.1 Datos

El *dataset* trabajado consta de 600 imágenes donde son 3 clases y cada una de ellas posee 200 imágenes. Estas imágenes se almacenan en una base de datos no relacional conocida como Firebase.

Llevar un registro de las aves será la principal herramienta para que el aplicativo pueda consultarla y saber si realmente es un animal en peligro de extinción (peligro crítico CR, en peligro EN, vulnerable VU) (véase la figura 8).

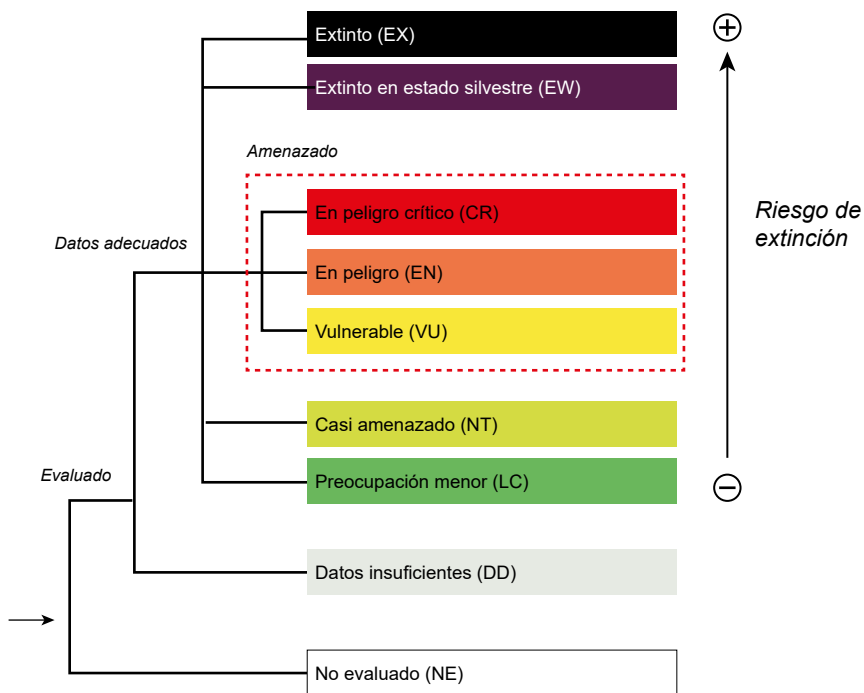


Figura 8. Categorías de amenaza según UICN

Fuente: Guía de identificación y cuidados iniciales de animales silvestres decomisados o hallados en abandono, Serfor (2016)

Luego de haber registrado a los animales haremos uso de las redes neuronales convolucionales para lograr una correcta identificación mediante el uso de imágenes.

## 4. ARQUITECTURA DE LA CNN (VGGNET16)

La arquitectura de la VGG16 consiste en un *InputLayer*, el cual acepta imágenes de 96 x 96 píxeles con una profundidad de 3, las cuales son procesadas por una capa de convolución en 2D que posee 32 filtros con núcleo de 3 x 3. Cada *pool* es de 3 x 3 que permite redimensionar las imágenes ingresadas de 96 x 96 a 32 x 32. Luego se realiza la activación y se va aumentando el tamaño de los filtros de 32 a 64. Se debe considerar que mientras más profunda es la red, menor serán las dimensiones del volumen y más filtros para aprender. Por consiguiente, se dimensiona el tamaño máximo de la agrupación de 3 x 3 a 2 x 2 para que las dimensiones no se reduzcan rápidamente. Estos pasos se realizan hasta llegar a un tamaño del filtro de 128 píxeles. Finalmente, en la capa totalmente conectada se realiza una activación lineal corregida y normalizada (Rosebrock, 2018).

### 4.1 Fases de la implementación de la CNN

a) *Creación y configuración de la CNN*

En esta primera fase, se realiza la creación de la clase *SmallerVGGnet* y se agregan las capas respectivas al modelo (Rosebrock, 2018).

b) *Entrenamiento de la CNN*

En esta segunda fase, se implementa el *script* para el entrenamiento de la red. Una vez implementada, se puede entrenar la CNN con Keras en base al entrenamiento deseado (Rosebrock, 2018).

c) *Clasificando imágenes con nuestra CNN y Keras*

En esta tercera fase, la red neuronal está totalmente entrenada y permite clasificar imágenes basándose de una secuencia de comandos (Rosebrock, 2018).

d) *Visualización de resultados*

En esta última fase, se analiza si la CNN ha clasificado correctamente la imagen del animal.

## 5. IMPLEMENTACIÓN

### 5.1 Descripción del *dataset*

Se ha construido un *dataset* de 600 imágenes repartidas en 3 clases de aves. Estas clases de aves fueron *Brotogeris\_versicolorurus*, *Carduelis\_magellanica* y *Sicalis\_flaveola*.

## 5.2 Data de validación y data para testing

Se utilizaron el 70 % de las imágenes para la validación y el 30 % para el *testing*. Esto se logró hacer con la función `sklearn.model_selection.train_test_split` que permite dividir un *dataset* en dos bloques, típicamente bloques destinados al entrenamiento y validación del modelo (llamemos a estos bloques “bloque de entrenamiento” y “bloque de pruebas”) (InteractiveChaos making things simple, s. f.).

## 6. RESULTADOS

Para el entrenamiento se utilizó una MacBookPro Intel Core i7 de séptima generación, 16 Gb de RAM y 4 Gb de Radeon Pro 560.

### 6.1 Métrica utilizada para el análisis

#### 6.1.1 Gráfico de precisión y pérdida

En la siguiente figura 9 se detalla el proceso de entrenamiento de la red durante las 1000 épocas. En cada una de las épocas existe un valor de entrenamiento perdido (*train\_loss*), valor perdido (*val\_loss*), valor ganado (*val\_acc*) y entrenamiento ganado (*train\_acc*). Estos valores van cambiando y generando una gráfica de la cual se concluye que el entrenamiento de la CNN es correcto ya que el valor ganado de las características obtenidas de las imágenes de animales superó al valor perdido durante el entrenamiento en la época 200 en adelante.

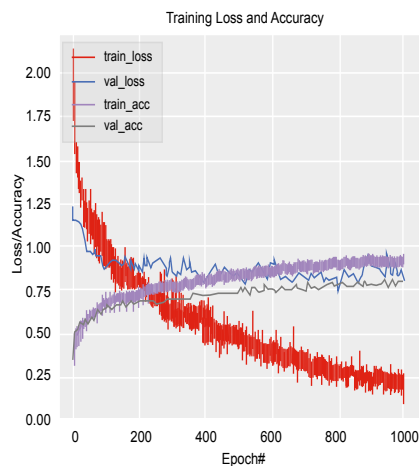


Figura 9. Gráfico del valor ganado y entrenamiento perdido, escenario B

### 6.1.2 Matriz

Una matriz de confusión resume el rendimiento de clasificación de un clasificador con respecto a algunos datos de prueba. Es una matriz bidimensional, indexado en una dimensión por la verdadera clase de un objeto y en el otro por la clase que el clasificador asigna. La figura 10 presenta un ejemplo de matriz de confusión para una clasificación de tres clases de aves (*Brotegeris\_versicolurus*, *Carduelis\_magellanica*, *Sicalis\_flaveola*). La primera fila de la matriz indica que 58 objetos pertenecen a *Brotegeris\_versicolurus* y que 39 fueron clasificados correctamente como pertenecientes a *Brotegeris\_versicolurus*, 9 clasificados erróneamente como pertenecientes a *Carduelis\_magellanica*; y 10 como pertenecientes a *Sicalis\_flaveola*.

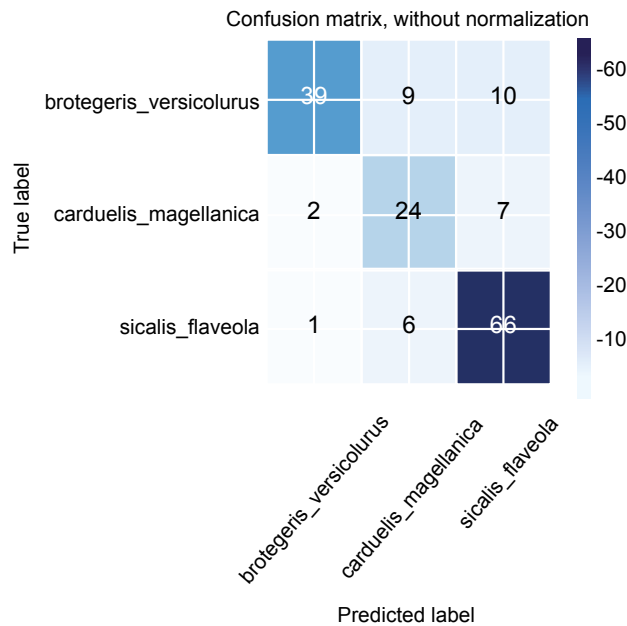


Figura 10. Matriz de confusión  
Elaboración propia

#### 6.1.2.1 Interpretación de las métricas de la matriz de confusión

Para la clase *Brotegeris\_versicolurus* se obtuvo una precisión de un 93 %, un *recall* de 67 % y un *f1-score* de 78 %; en la clase *Carduelis\_maggellanica*, una precisión de un 62 % , un *recall* de 73 % y un *f1-score* de 67 %; en la clase 2, una precisión de un 80 % , un *recall* de 90 % y un *f1-score* de 85 %. Se concluye que el modelo y las predicciones serán correctas (véase la figura 11).

	precision	recall	f1-score	support
0	0.93	0.67	0.78	58
1	0.62	0.73	0.67	33
2	0.80	0.80	0.85	73
avg/total	0.81	0.79	0.79	164

Figura 11. Resumen de los indicadores por cada clase escenario A  
Elaboración propia

## 7. DISCUSIÓN

La precisión promedio de 81 % demuestra que la identificación de imágenes es una característica propia de las redes neuronales convolucionales. Los hallazgos sobre la identificación de objetos abrieron las puertas a dicha investigación. El hecho que desarrollemos una aplicación que tenga la capacidad de identificar a un ave es beneficioso porque agilizará el proceso de intervención policial de animales. Asimismo, se construyó la base de datos no relacional que llevará los registros de las aves en tiempo real. La razón de la elección de una base de datos no relacional se debe a que no es posible llegar a la normalización de la *data* por la diversidad de información. Esta investigación brindará las bases para la posible investigación de clasificación de animales de todas las especies agrupadas por subfamilias y familias.

A medida que el campo de la investigación sea aplicado a la fauna silvestre permitirá lograr nuevos avances y experimentos basados en hipótesis. A nivel mundial existe una creciente necesidad por brindar aportes a la fauna silvestre debido a que el tráfico ilegal de animales ocupa el tercer comercio ilegal.

Aunque identificar correctamente un ave es un primer paso crítico, se debe considerar que las especies trabajadas poseen una taxonomía completa. Por ende, como trabajo futuro se ha proyectado el uso de redes generativas antagónicas (GAN) con el fin de poder reconstruir alguna extremidad o parte de la especie que no se encuentre completa.

Finalmente, nuestro enfoque tiene como objetivo aplicar técnicas de inteligencia artificial a la fauna y de ese modo mitigar el tráfico ilegal de animales.

## 8. CONCLUSIÓN

Se construyó un *dataset* de imágenes de aves y se evaluaron diversas métricas como la precisión, *recall*, *f1-score*, *support* con el fin de identificar automáticamente tipos de ave. Al utilizar nuestro conjunto de datos se proporcionaron resultados prácticos y análisis de desempeño del reconocimiento a lo largo del entrenamiento. Los resultados experimentales demostraron la posibilidad de utilizar el reconocimiento de imágenes para la clasificación de un solo dominio de especie (aves).

Actualmente, se está trabajando en la expansión de reconocimiento de familias o subfamilias de diversas especies. Finalmente, nos gustaría enfatizar que nuestra aplicación pueda mitigar el tráfico ilegal de animales.

## REFERENCIAS

- Barletta, F., Pereira, M., Robert, V., y Yoguel, G. (2013). Argentina: dinámica reciente del sector de *software* y servicios informáticos. *Revista de la CEPAL*, (110), 137-155. Recuperado de <http://www.cepal.org/publicaciones/xml/1/50511/RVE110Yoqueletal.pdf>
- Carrión, A. (2017). Tráfico de fauna silvestre en estado de emergencia. *El Comercio*. Recuperado de <https://elcomercio.pe/wuf/noticias/trafico-fauna-silvestre-emergencia-402805-noticia/>
- InteractiveChaos making things simple. (s. f.). InteractiveChaos making things simple. Recuperado de <https://www.interactivechaos.com/python/function/traintestsplit>
- Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., y Fei-Fei, L. (2014). Large-scale video classification with convolutional neural networks. *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 1725-1732.
- López Tarabochia, M. (2016). Perú: Rescatan 25 aves silvestres que iban a ser comercializadas ilegalmente en Europa. *Revista Mongabay*. Recuperado de <https://es.mongabay.com/2016/12/peru-aves-trafico/>
- Marques, A. C., Raimundo, M. M., Cavalheiro, E. M. B., Salles, L. F. P., Lyra, C., y von Zuben, F. J. (2018). Ant genera identification using an ensemble of convolutional neural networks. *PLoS ONE*, 13(1).
- Motta, D., Santos, A. Á. B., Winkler, I., Machado, B. A. S., Pereira, D. A. D. I., ... Badaró, R. B. (2019). Application of convolutional neural networks for classification of adult mosquitoes in the field. *PLOS ONE*, 14 (1), 1-18.
- Musso, D., Rodriguez-Morales, A. J., Levi, J. E., Cao-Lormeau, V.-M., y Gubler, D. J. (2018). Unexpected outbreaks of arbovirus infections: lessons learned from the Pacific and tropical America. *The Lancet Infectious Diseases*, 18(11), 355-361. doi:10.1016/S1473-3099(18)30269-X

- Oquab, M., Bottou, L., Laptev, I., y Sivic, J. (2014). Learning and transferring midlevel image representations using convolutional neural networks. *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 1717-1724.
- Quevans, N., Falcón, N. y Elías, R. (2014). Fauna silvestre y productos derivados decomisados durante el período 2000-2007, Lima-Perú. *Salud y Tecnología Veterinaria*, 1, 14-18. doi:10.20453/stv.2013.105
- Rosebrock, A. (2018). Keras y redes neuronales convolucionales (CNN). *Pyimagesearch*. Recuperado de <https://www.pyimagesearch.com/2018/04/16/keras-and-convolutional-neural-networks-cnns/>
- Rossius, S. (2013). Reconocimiento de objetos mediante WebCam en tiempo real (Doctoral dissertation).
- Serfor. (2017) Estrategia Nacional para reducir el tráfico ilegal de fauna silvestre en el Perú 2017-2027 y su plan de acción 2017-2022. Recuperado de <https://www.serfor.gob.pe/wp-content/uploads/2017/09/PRESENTACION-DE-LA-ESTRATEGIA-NACIONAL-PARA-REDUCIR-EL-TRAFICO-ILEGAL-DE-FAUNA-SILVESTRE.pdf>
- Serfor. (2016). *Guía Identificación y cuidados iniciales de animales silvestres decomisados o hallados en abandono*. Recuperado de <https://www.serfor.gob.pe/wp-content/uploads/2016/08/Guia-de-identificacion-final12ago.pdf>
- S. R. (2013). *Reconocimiento de objetos mediante WebCam en tiempo real*. Universidad Politécnica de Valencia.
- Tobías, L., Ducournau, A., Rousseau, F., Mercier, y G., Fablet, R. (2017). Convolutional Neural Networks for object recognition on mobile devices: A case study. *2016 23rd International Conference on Pattern Recognition (ICPR)*, 3519-3524.
- Zeiler, M. D., y Fergus, R. (2014). Visualizing and understanding convolutional networks. *European conference on computer vision*, 8689, 818-833.





# Aprender a programar sin el uso de la computadora

Marjorie Tineo-Tineo  
alizmttineo@gmail.com / Universidad Continental, Perú

Recepción: 16-6-2019 / Aceptación: 8-8-2019

RESUMEN. Este proyecto tiene como objetivo lograr que los niños y niñas puedan programar sin utilizar una computadora con la metodología ágil en el distrito de Chuschi, provincia de Cangallo, perteneciente al departamento de Ayacucho. Las escuelas existentes en el distrito cuentan con un aula de informática donde el docente a cargo enseña a los estudiantes la manipulación del Word y PowerPoint, además, las niñas y los niños presentan dificultad en el aprendizaje de las matemáticas. Por ello, este proyecto pretende enseñar a los estudiantes lo que podrían hacer con la programación antes de estudiarla mediante la integración de las áreas de matemáticas y arte. Las actividades del plan de trabajo fueron divididas en dos partes, la primera parte desarrolla el pensamiento computacional; la segunda, la programación lógica. Se utilizaron tres currículos de los programas de CS Unplugged, codeSpark Unplugged y Code Unplugged; el primero enseña conceptos de informática, el segundo brinda una noción de la programación y el tercero imparte conocimientos de programación. En todas las actividades se emplearon conocimientos, uso de estrategias, técnicas lúdicas de aprendizaje y enseñanza.

PALABRAS CLAVE: metodología ágil, pensamiento computacional, programación lógica, programación

## Learning to Program Without a Computer

ABSTRACT. This project aims to achieve boys and girls to be able to program without a computer with the agile methodology in the district of Chuschi, province of Cangallo, department of Ayacucho. The district's schools have a computer classroom where students learn programs such as Word and PowerPoint. In addition, girls and boys have difficulty in learning math. Therefore, this project aimed to teach students what they could do with programming before studying it through the integration of math and art. The curriculum's activities were divided into two parts: the first one developed the computational thinking and the second one, the logical programming. CS, codeSpark and Code unplugged activities were used: the first one teaches computer concepts, the second one provides programming basics and the third one teaches programming skills. In all the activities, knowledge, strategies, and playful techniques of learning and teaching were used.

KEYWORDS: agile methodology, computational thinking, logical programming, programming

## 1. INTRODUCCIÓN

El presente proyecto social se enfoca en aprender a programar cuyo objetivo principal es lograr que el niño(a) logre programar sin utilizar una computadora.

En PIXIE (11 de febrero del 2016) señalan, “que estudiar programación se ha vuelto una herramienta clave para nosotros que nos ayuda a enfrentar este nuevo mundo laboral tecnológico. En el futuro no solo se demandará de programadores, sino de profesionales que dominen el pensamiento computacional y sepan transmitirlo a los campos de la medicina, la biología, el diseño y las artes. Cabe recalcar que enseñar programación a edades tempranas derriba preconceptos de género y anima a las niñas a pensar también en la posibilidad de estudiar carreras tecnológicas” (p. 1).

El programar enseña a pensar desarrollando la habilidad de resolución de problemas. Además, se desarrolla la creatividad, la lógica y el razonamiento. La programación refuerza los conocimientos matemáticos y el trabajo en equipo.

Según Andrés Karp, presidente de ADWE (Asociación de Desarrolladores Web de España), la programación y el superar los constantes desafíos que ello conlleva induce a los niños a utilizar cada vez más su creatividad para avanzar en el desarrollo de lo que puedan estar programando. No hay una única solución para llegar a un camino en la programación, el ejercicio de la creatividad les permitirá alcanzar las metas que se hayan propuesto al momento de programar. Los niños no solo van a interactuar en un entorno tecnológico, sino que van a crear. Pueden construir sus propias herramientas, asociar su creatividad con el desarrollo de esas ideas transformándolas en código.

Además, considera que aprender a programar a una edad temprana favorece, en gran medida, la creatividad de aumentar la capacidad de resolución de problemas. Pablo Espeso, ingeniero informático y coordinador del Club de Jóvenes Programadores de la Universidad de Valladolid, afirma que programar es pura creatividad, siempre y cuando se sepa llevar a cabo.

En la web de Requetic (30 de junio del 2018), comentó que la programación ayuda a los niños a utilizar la creatividad, incluso, plantear la resolución de problemas de razonamiento lógico o matemático. En el blog de Silabuz (20 de julio del 2018), asegura que muchos estudios hechos en niños demuestran que el ejercicio de la programación en edades tempranas ayuda en la mejora de la capacidad de cálculo y lógica. Estos dos atributos están integrados en la acción de programar.

Para la ejecución del proyecto se preparó un manual de actividades que fue una recopilación de los currículos de CS Unplugged, codeSpark Unplugged y Code.org. Al momento de impartir la enseñanza de las lecciones y la realización de las actividades se utilizó la metodología ágil.

El proyecto inició el 9 de enero hasta el 15 de febrero 2019, los días miércoles, jueves y viernes con un total de 41 horas de estudio en una duración de seis semanas. En la quinta y

sexta semana se citaron a los estudiantes a asistir en la tarde de forma voluntaria para aplicar lo aprendido en las computadoras utilizando *software* proporcionado de la Hora del Código de diferentes países.

En un comienzo los beneficiarios fueron diecisiete niños y niñas, sin embargo, el proyecto culminó con un total de once niños y niñas de siete a nueve años.

Al finalizar se realizó una encuesta a los representantes de los beneficiarios y se solicitó a cada uno de ellos que grabaran en un video entrevista a los beneficiarios para saber la opinión de viva voz de cada uno de ellos.

## 2. METODOLOGÍA

Se utilizó la metodología ágil, con el objetivo de que los beneficiarios lideren su propio aprendizaje, mediante dinámicas estimulantes que les permitan crecer en autonomía, aprender a trabajar en equipo y buscar siempre la mejora continua. Hubo un total de tres Sprint, cada uno de ellos tuvo una duración de dos semanas. Se utilizó un tablero de gestión visual conformado por el nombre del equipo, columnas y casillas del tablero. Las columnas contenían las siguientes variables: temas, tareas, en curso, por validar y hechas. En las casillas se incluían: *stop*, impedimento y retrospectiva.

Por ejemplo, en la actividad Magia de voltear cartas - Detección y corrección de errores de CS *unplugged*, el objetivo fue demostrar cómo detectar los datos cuando se han corrompido y cómo corregirlos. Esta actividad se desarrolló en dos días dividido en dos partes.

### 2.1 Primera parte, 16 de enero

Para comenzar se informó que se enseñaría un truco de magia, pero antes de aprenderlo, se les preguntó que a cuántos de ellos al ver una película, de repente vieron que se congeló la imagen o empezó a verse en cuadraditos, al tiempo que sonaba un ruido como de “rayado”; acto seguido se les explicó que eso sucedía porque el DVD no funcionaba correctamente, para lo cual se tenía que sacar el DVD del reproductor, limpiarlo y regresarlo al reproductor para que funcionara, ahora sí, correctamente.

Otro ejemplo que se compartió con ellos fue cuestionarlos sobre el uso de la computadora por parte de sus profesores, quienes suelen utilizarla para escribir informes escolares, entre ellos, sus calificaciones. Se les compartió el siguiente relato:

Supongamos que su profesor de escuela está redactando las notas de cada uno de ustedes y de repente la información almacenada se cambia sin que nadie lo sepa. Por ejemplo, sobre la calificación de Pedro, el profesor la redactó con la nota “A”, guarda el documento y se va al almorzar dejando la computadora. De pronto la nota “A” cambia a “C” sin que el profesor lo sepa. A su regreso, solo

le basta imprimir el documento y acto seguido entrega la boleta de notas en clase. Pedro se da con la sorpresa que tiene todas las calificaciones con “C”, se lo hace saber al profesor, quien lo ve con asombro y se pregunta: ¿cómo pudo pasar?

Sobre este relato, se les preguntó a los estudiantes, si el profesor para la próxima ocasión escribiría la boleta de notas en la computadora, sí o no, todos respondieron, no, ¿por qué?

Los alumnos respondieron sobre la posibilidad de volverlo a hacer mal, por eso el profesor lo escribiría en papel. Se les explicó que los ordenadores cuando detectan ese tipo de errores, lo que pasó con Pedro, lo corrigen devolviéndose a como estaba antes del cambio. Fue así como se les explicó que en la actividad verían un truco de magia donde se analizaría cómo las computadoras corrigen esos errores automáticamente.

Se solicitó un voluntario a quien se le entregaron unas tarjetas, se le pidió que colocara una fila de cinco cuadraditos con algunas que mostraran las negras y la otra blanca. La instrucción, después, fue que agregara otra fila y así sucesivamente hasta que se tuviera una cuadrícula de 5 x 5. A continuación, se le dijo: “Voy a hacerlo un poco más difícil agregando otra fila y otra columna”.

Después se pidió otro voluntario, a quien se le explicó que debía voltear solo un cuadrado, fuese negro o blanco, en el momento en que la instructora se volteara, porque iba a mirar y averiguar cual se había cambiado, para ello, cuando se hiciera el cambio, todos debían decir: “Listo”, como señal para darse vuelta y adivinar. Al girarse, detectó cual fue, la volteó y preguntó: “¿Acerté, sí o no?”. Todos respondieron que sí.

Se volvieron a hacer cuatro ejemplos más, para que se dieran cuenta del patrón, muy pocos lo identificaron, sin embargo; se explicó cuál era el patrón que obedecía: el número par; es la suma que todas las filas y columnas tiene en total los cuadrillos negros, si se cambia cualquiera de una de ellas al voltearla, esa fila y esa columna pasa a ser impar, por tanto, se ve la intersección de esa ficha con la columna y la fila impar, se voltear a su posición original detectando así el cambio con este truco.

A continuación, se les indicó que escriban en un *post-it*, “magia de voltear cartas” y lo colocaran en la columna de temas; luego escribir el número de la actividad “1” en la de tareas, la cual se iba a desarrollar en equipos (había cuatro grupos).

Se entregó la hoja de actividades a cada equipo que consistió en recortar los cuadraditos y formar el cuadrado cumpliendo el patrón de paridad. Cada uno tenía que adivinar entre ellos qué cuadrado volteaba su compañero, para ello usaron hojas en blanco para los costados de las filas y columnas, sumando cuantos cuadrados negros había en cada uno, identificando así quien era impar en ambos, de esta manera hallaban la fila y la columna que se modificaba.

Antes de culminar la clase, se rindió una evaluación grupal a cada uno de sus miembros para saber si habían logrado comprender el patrón, en donde alguien cambiaba un cuadrado y se adivinaba cual era. Se concluyó que ese mismo patrón identificado era el mismo que las computadoras hacían al detectar y corregir los errores.

## 2.2 Segunda parte, 17 de enero

Siguiendo con la idea de unir los relatos cotidianos como puente para explicar la programación, se les preguntó si alguna vez habían visto en una tienda un aparato que al señalar al código de barras del producto, sonaba un ruido y acto seguido en la pantalla de la máquina que va registrando las compras se ven los números, luego sigue así hasta terminar de escanear todos los productos, que papá o mamá está comprando. Así, por último, la cajera dice el valor a pagar de la compra, valor al que llegó con solo escanear el código de barras de cada producto. Sin embargo, a veces cuando el lector del código de barras no logra identificar el producto, la cajera vuelve a repetir el mismo proceso hasta que se reconozca el producto y, por ende, el monto.

Entonces, ¿cómo logra reconocer la máquina cuándo está mal o bien un producto? Para saberlo, se calculará el último dígito de un código de barras de trece dígitos, con una actividad para identificar dígitos de verificación del código del producto.

Para la demostración se usaron empaques de galletas que tienen el código de barras. Cada niño dictaba todos los dígitos excepto el último, entonces, se calculaba cuál podría ser el último dígito y verificábamos si estaba bien o no al visualizarlo en el envoltorio. Para sorpresa, se lograron identificar. Luego se hizo el mismo ejercicio tres veces más, esta vez usando códigos de barras de los libros.

Después de esta explicación se entregaron las hojas de actividades en las cuales consistía hallar los últimos dígitos de tres productos. Estas actividades se desarrollaron de manera grupal, donde se observó cómo hallaban el último dígito según el método enseñado. Antes de finalizar la clase, se rindió una evaluación sobre la actividad realizada.

Cuando finalizaba cada sprint se otorgaron premios a todos los equipos según las siguientes categorías:

- Primera categoría: ¿Quién aprendió más?
- Segunda categoría: ¿Quién mejoró más?
- Tercera categoría: ¿Quién trabajó más?

También se utilizaron las categorías mencionadas en la premiación individual en el cierre del proyecto. En una categoría hubo más de un equipo y estudiante premiado. Con la metodología ágil se llegó a premiar a todos los estudiantes sin excepción, a diferencia del modelo tradicional que solo otorga a los tres primeros puestos.

## 3. RESULTADOS

Al no hacer uso de la computadora, este proyecto no se alejó de los conceptos de la informática. Las actividades se basaron en conceptos matemáticos, la tecnología en el conocimiento y

entendimiento de cómo trabajan las computadoras. Las temáticas desarrolladas y aprendidas por las niñas y los niños fueron sobre los números binarios, la representación de imágenes, la detección y corrección de errores, los lenguajes de programación, el diseño de la interfaz humana, la prueba de Turing, la introducción a los fundamentos de la programación con los Foos, la construcción de Marble Run, los algoritmos de la vida real, las secuencias con mapas felices, programándose entre sí mismos con muévete, muévete; la seguridad en el entorno e internet, entendiendo con bucles, control de eventos con Jr. y el construir una figura base para superar el reto de sostener dos libros medianos por 10 segundos.

Se logró aplicar lo aprendido en los juegos *offline*: LightBot, Condorito, Star Wars, Minecraft, Pilas y Bloques, pertenecientes a la Hora del Código, en donde casi no tuvo que intervenir la investigadora porque los estudiantes lograban entender perfectamente el desafío. Solo intervino en los niveles de dificultad en lo que se hacía más uso de condicionales y bucles porque se complicaban en tratar de colocar los comandos correctos; incluso, a veces usaron comandos innecesarios, dificultándoles el tratar de entender lo que habían hecho al ejecutar su programa, no logrando cumplir con el reto y, por ende, no llegaban a pasar al siguiente nivel. Sin embargo, cuando se les orientó en donde estaban fallando, sin corregir completamente el programa, lograron captar qué estaba mal y entre ellos mismos hallaron el error de su código.

Por ejemplo, en el juego LightBot, cuando los alumnos alcanzaban el nivel 2-4, no solo había un bucle (P1), había dos bucles (P1 y P2), es aquí donde hubo dificultad, porque redundaron pasos en los bucles P1 y P2, necesitando pasos adicionales para lograr la meta. La investigadora intervino con explicaciones al momento de reproducir el juego, en donde había pasos que el robot redundaba, como caminar dos veces el mismo bloque, entonces, comentó que el robot debía recorrer el camino más corto para llegar a los bloques azules; fue así que tres alumnos que trabajaban en una computadora lograron captar el error de su secuencia de pasos que establecieron, modificando así lo que era necesario, de esta manera consiguieron cumplir el reto pasando al siguiente nivel.

*Aclaración.* Cuando se realizó la actividad desenchufada de bucles, los alumnos desarrollaron satisfactoriamente obteniendo buenos resultados, la dificultad se presentó en los juegos *offline* de la Hora del código.

Hubo momentos en que se frustraron y se divirtieron en el transcurso de su aprendizaje de aprender a programar sin computadora y también cuando lo aplicaban en el ordenador. Para la frustración en las actividades desenchufadas, las niñas y los niños tuvieron un entrenamiento en la construcción del Marble Run y del Glitch. En la lectura de Stevie y el gran proyecto se brindaron recomendaciones mediante la historia de la ardilla de como manejar la frustración y lo positivo que se puede sacar de ello. Siempre se recordó esta historia a los estudiantes en el transcurso del desarrollo del plan de trabajo, poniendo así en práctica su persistencia en cumplir el desafío de las actividades.

El mayor impacto que tuvo el trabajar el proyecto con los niños y las niñas, fue en lo personal, pues este tipo de trabajo brinda satisfacción al interactuar con ellos. Además, el ver como se involucraron activamente en las habilidades de comunicación, en la resolución de problemas, desarrollando su creatividad y mejorando sus habilidades blandas. Esto fue más visible el último día de clases, donde ya no tenían que trabajar en equipos formados sino individualmente, sin embargo todos se apoyaron mutuamente porque ya habían trabajado anteriormente en diferentes grupos con todos sus compañeros.

Según los resultados de la encuesta, el 100 % de los representantes de los beneficiarios conocieron los resultados del proyecto y afirmaron que atendieron las necesidades de su población. Según propias palabras en las sugerencias de la encuesta de satisfacción del representante de Luis Fernando: “La población estudiantil requiere de este tipo de apoyo ya que fortalece el conocimiento en esta edad temprana”. Adicional a ello, la Superiora de las Hnas. de la Caridad de Leavenworth solicitó la posibilidad de repetirse este tipo de proyectos en el distrito de Chuschi.

Se cumplió con el plan de trabajo establecido desarrollándose íntegramente las catorce actividades estipuladas.

En la siguiente hoja se muestra la estadística del resultado de la encuesta de satisfacción a los representantes de los beneficiarios.

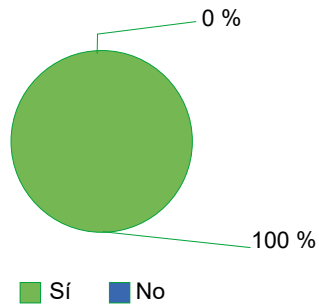
### Resultados de la encuesta de satisfacción a los grupos de interés

Fecha: 15/02/2019  
 Nombre: Annie Liz Marjorie Tineo Tineo  
 Proyección Social: Profesión para el desarrollo

#### 1. *¿Conoce los resultados del proyecto?*

Ord	Opción de respuesta	Frecuencia	%
1	Sí	14	100
2	No	0	0
Total		14	100

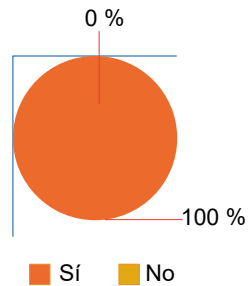
**Frecuencia de los resultados del proyecto**



2. *¿El proyecto atendió a las necesidades de su población?*

Ord	Opción de respuesta	Frecuencia	%
1	Sí	14	100
2	No	0	0
Total		14	100

**Frecuencia de la atención a las necesidades de la población**

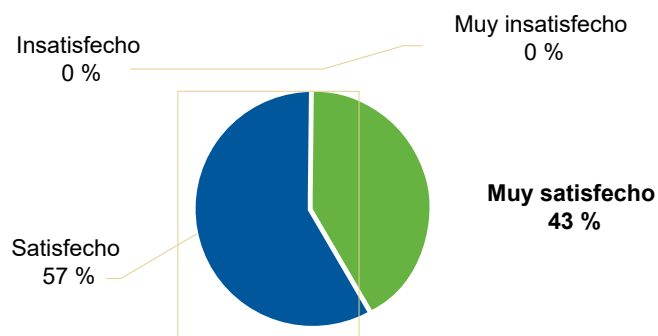




3. *¿Qué tan satisfecho o insatisfecho está usted con el proyecto?*

Ord	Opción de respuesta	Frecuencia	%
1	Muy satisfecho	6	43
2	Satisfecho	8	57
3	Insatisfecho	0	0
4	Muy insatisfecho	0	0
Total		14	100

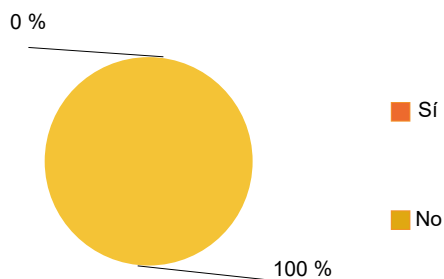
**Frecuencia de satisfacción con el proyecto**



4. *¿Ha recibido algún tipo de ayuda social de otra institución educativa superior?*

Ord	Opción de respuesta	Frecuencia	%
1	Sí	14	100
2	No	0	0
Total		14	100

**Frecuencia de haber recibido algún tipo de ayuda social de otra IES**



#### 4. CONCLUSIONES

Hasta la fecha no había ningún proyecto social educativo ejecutado en la zona del Perú por parte de las universidades u organizaciones.

- El motivo por el cual se escogió Chuschi, como piloto en la ejecución de este proyecto, más allá de lo que vivieron en tiempo del terrorismo, la intención fue demostrar que si se brinda la oportunidad a las niñas y los niños de Chuschi que se da en la ciudad de Lima en el aprendizaje de la programación, se mostrara que no hay diferencias gigantescas entre los las niñas y los niños de Chuschi y las niñas y los niños de Lima, ambos tienen las mismas capacidades para aprender. Muchos tienen hasta ahora la mentalidad de que las niñas y los niños de zonas rurales no tienen suficiente capacidad de aprender y/o no son tan buenos estudiantes, lo cual es falso, se ha visto mediante las evaluaciones de las niñas y los niños de las diferentes edades un gran avance en el aprendizaje y captación de los contenidos aprendidos en las lecciones realizadas.
- Cabe destacar que las lecciones impartidas a los estudiantes del proyecto ya habían sido replicadas en Nueva Zelanda, Estados Unidos, Canadá, Suecia, Australia, China, Corea, Taiwán, México, Italia y Canadá. En el Perú, con la ejecución de este proyecto viene a ser la primera réplica a nivel nacional.
- Se recomienda continuar con este tipo de iniciativas en beneficio de los niños y niñas de las zonas rurales.
- Se recomienda una visita del programa Enciende Cultura al distrito de Chuschi.

#### REFERENCIAS

- Albaladejo, G., y Albaladejo, X. (2018). *Agilizando las aulas - Guía para implementar la metodología ágil en clase*. Recuperado de <https://clasesagiles.files.wordpress.com/2018/01/guia-metodologia-agil-en-clase-v1-01.pdf>
- Bell, T., H., I. y Fellows, W. (2008). *Un programa de extensión para niños de escuela primaria* (pp. 7-28, 37-42, 107-111). Recuperado de <https://classic.csunplugged.org/wp-content/uploads/2014/12/unpluggedTeachersDec2008-Spanish-master-ar12182008.pdf>
- Bell, T., H., I. y Fellows, W. (2015). *An enrichment and extension programme for primary-aged students* (pp. 15, 17, 209-243). Recuperado de [https://classic.csunplugged.org/wp-content/uploads/2015/03/CSUnplugged\\_OS\\_2015\\_v3.1.pdf](https://classic.csunplugged.org/wp-content/uploads/2015/03/CSUnplugged_OS_2015_v3.1.pdf)
- Cinco Días. (2016). *¿Por qué es importante aprender a programar?* Recuperado de [https://cincodias.elpais.com/cincodias/2016/02/11/tecnologia/1455182591\\_761653.html](https://cincodias.elpais.com/cincodias/2016/02/11/tecnologia/1455182591_761653.html)

- Code.org. (2018a). *Curriculum for Course 1-Code.org*. Recuperado de <https://code.org/curriculum/course1>
- Code.org. (2018b). *Curriculum for Course 2-Code.org*. Recuperado de <https://code.org/curriculum/course2>
- Code.org. (2018c). *CS Fundamentals 2018-Code.org*. Recuperado de <https://curriculum.code.org/csf-18/coursea/>
- Code.org. (2018d). *CS Fundamentals 2018-Code.org*. Recuperado de <https://curriculum.code.org/csf-18/courseb/>
- CodeSpark Academy. (s. f.). *The Foos Activity Book* (pp. 1-16). Recuperado de [https://thefooscom.s3.amazonaws.com/cms/admin-uploads/codeSpark\\_Activity\\_Book.pdf](https://thefooscom.s3.amazonaws.com/cms/admin-uploads/codeSpark_Activity_Book.pdf)
- Compromiso Disney. (s. f.). *Amigos conectados Ficha 1 - Ficha 7*. Recuperado de <http://compromiso.disneylatino.com/creatividad-innovacion>
- Csunplugged.org. (2018a). *Topics - CS Unplugged*. Recuperado de <https://csunplugged.org/en/topics/>
- Csunplugged.org. (2018b). *Para imprimir - CS Unplugged*. Recuperado de <https://csunplugged.org/es/resources/>
- PIXIE. (11 de febrero del 2016) *¿Por qué es importante aprender a programar?* Recuperado de [https://cincodias.elpais.com/cincodias/2016/02/11/tecnologia/1455182591\\_761653.html](https://cincodias.elpais.com/cincodias/2016/02/11/tecnologia/1455182591_761653.html)
- Requetic. (30 de junio del 2018). *Scratch JR: programación para niños a partir de 5 años*. Recuperado de <http://www.requetetic.com/blog/scratch-jr-programacion-para-ninos/>
- Silabuz. (20 de julio del 2018). *7 beneficios del aprendizaje de programación en los niños*. Recuperado de <http://blog.silabuz.com/2018/07/20/7-beneficios-del-aprendizaje-de-programacion-en-los-ninos/>



# Propuesta de un patrón de arquitecturas de *software* para la interoperabilidad en dispositivos en la capa al borde de un ecosistema IoT

Juan Moreno-Motta

juan.moreno2@unmsm.edu.pe / Escuela de Ingeniería de Sistemas e Informática,  
Universidad Nacional Mayor de San Marcos, Lima, Perú

Felipe Moreno-Vera

felipe.moreno@ucsp.edu.pe / Departamento de Ciencia de la Computación, Universidad  
Católica San Pablo, Arequipa, Perú

Frank Moreno-Vera

fmorenov@uni.pe / Escuela de Ingeniería Electrónica, Universidad Nacional de Ingeniería,  
Lima, Perú

Recepción: 31-5-2019 / Aceptación: 9-7-2019

RESUMEN. En los últimos años el Internet de las cosas ha generado una disrupción en el ecosistema de soluciones para usuarios finales con aplicaciones en la salud, agroindustria, medio ambiente y otras soluciones heterogéneas. Cada una de estas tienen su propio formato para verbalizar los datos proporcionados por los diferentes servicios SOAP, RESTful, REST-LD en formato JSON o XML que pueden ser entendidos por personas. Estos dispositivos se denominan “heterogéneos” porque provienen de diferentes proveedores de manufactura, diferentes lenguajes de programación como C, C++, Lua, Python, JavaScript, etc. Este documento se enfoca sobre dispositivos restringidos, es decir, dispositivos con poca capacidad de procesamiento y memoria con diferentes cadenas de datos sensados. Esta investigación propone un patrón de arquitectura de *software* para la interoperabilidad entre los dispositivos al borde de un ecosistema IoT y entre ecosistemas entre sí.

PALABRAS CLAVE: interoperabilidad, IoT, dispositivos restringidos, microservicios, semántica web, *encoder*, *decoder*, *fog computing*, *edge computing*, REST API, *software architecture*, *IoT Ecosystem*

## Proposal for a Software Architecture Pattern for the Interoperability Between Devices on the Edge of an IoT Ecosystem Layer

**ABSTRACT.** In recent years, the Internet of Things has generated a disruption in the solution ecosystem for end users with applications in health, agro-industry, environment and other heterogeneous solutions. Each of these solutions has its own format to verbalize the data provided by different web services, including SOAP, RESTful and REST-LD, in JSON or XML format, so that said data can be understood by people. These devices are called “heterogeneous” because they come from different manufacturers and use different programming languages such as C, C++, Lua, Python, Javascript, etc. This document focuses on constrained devices, that is, devices with little processing capacity and memory, and different chains of acquired data. This research proposes a software architecture pattern for interoperability between devices on the edge of an IoT ecosystem and between ecosystems.

**KEYWORDS:** interoperability, IoT, constrained devices, microservices, semantic web, encoder, decoder, fog computing, edge computing, REST API, software architecture, IoT ecosystem

## 1. INTRODUCCIÓN

### 1.1 Revisión previa

El *Internet of Things* (de ahora en adelante, IoT) no es una tecnología nueva, es la convergencia de muchas tecnologías para soluciones como la salud, medio ambiente, manufactura, ahorro de energía, seguridad ciudadana. El IoT es un sistema de objetos físicos que se pueden descubrir, controlar o interactuar con dispositivos electrónicos que se comunican a través de varias interfaces de red y que, finalmente, se pueden conectar a Internet (Guinard y Trifa, 2016). Hasta hace poco, los proyectos de IoT se centraban principalmente en la creación de implementaciones a pequeña escala, cerradas y aisladas en las que los dispositivos no estaban diseñados para ser fácilmente accesibles o reprogramables (Guinard y Trifa, 2016).

En el año 2014, según en el estudio realizado por Oliver Klein (2017), se informó que en el año 2013 existían 1 billón de unidades de Smartphone y que para el año 2017 se esperaba 1.7 billones. En el mismo año, Jaimini (2017) indicó que en el año 2014 National Health Interview Survey realizó un estudio sobre el asma que afecta a los niños en EE. UU. lo cual conlleva a la pérdida de días de escuela y otros costos sociales. En ese año se estimó que para el año 2017 se incrementarían los dispositivos IoT a 15 mil millones y que el almacenamiento de los datos superaría 150 *exabytes*.

### 1.2 Actualidad

Talavera *et al.* (2017) realizan un trabajo de estado del arte sobre el número de implementaciones de IoT en los sectores de la agroindustria y el medio ambiente, tomando como referencia los resultados de otros, los clasificaron en identificación (3578), en selección (2652), en elegibles (720) y se tomaron finalmente 72 proyectos. Como resultado presentaron una arquitectura que trata de consolidar el ancho espectro de implementación (véase la figura 1). Uno de los retos que consideran en su trabajo de investigación es que una estandarización más sólida ayudaría a mejorar la compatibilidad entre diferentes proveedores y garantizar las medidas de seguridad más sólidas en todas las capas de IoT, desde los dispositivos de campo hasta los proveedores de la nube y las interfaces de usuario final. Este punto se refiere a la interoperabilidad entre las distintas capas de comunicación.

También en el año 2017, se desarrolló una plataforma para el control y monitoreo de la salud de personas dentro de sus hogares. Esta plataforma fue denominada SPHERE (Sensor Platform for HEalthcare in Residential Environment). Este proyecto determinó que existen nueve requerimientos que el IoT necesita cubrir y uno de ellos es la interoperabilidad. En el caso de dispositivos restringidos, es decir, con baja capacidad de memoria y procesamiento, Elsts, Oikonomou, Fafoutis y Piechocki (2017) consideran que los sistemas deben cumplir con los estándares y protocolos de IoT existentes de baja potencia para 1) ser susceptible de futuras extensiones con componentes de terceros; 2) reducir el tiempo de aprendizaje para el nuevo personal.

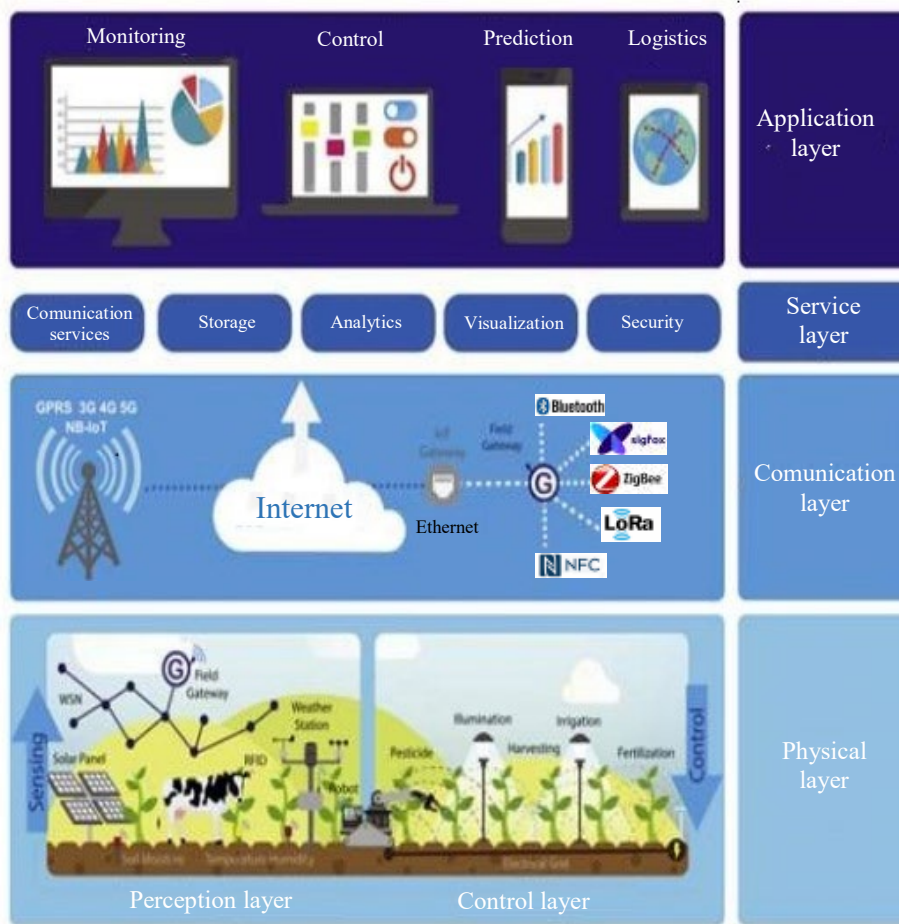


Figura 1. Propuesta de arquitectura IoT para aplicaciones agroindustriales y ambientales

Fuente: Datos tomados de Talavera *et al.* (2017)

Existen muchas soluciones de IoT. Cada implementación de IoT presenta retos de requerimientos no funcionales como son seguridad, almacenamiento, escalabilidad, interoperabilidad y energía. Alkhalil y Ramadan (2017) mencionan que los retos relativos a la procedencia de los datos, a pesar de las técnicas actuales, sigue siendo un desafío en su implementación y optimización. Por lo cual el reto de procedencia de datos se enlaza directamente con la interoperabilidad.

En la actualidad, la interoperabilidad es uno de los desafíos en las implementaciones de soluciones de IoT. Se realiza demasiado esfuerzo al integrar la información generada por los diferentes dispositivos en el dominio de salud, agroindustria y medio ambiente. Dentro de cada dominio existen diferentes sensores, para cada sensor existen diferentes marcas y fabricantes



(Pace *et al.*, 2017). Cada marca y fabricante trae consigo herramientas de desarrollo, con su propio lenguaje de programación. Los datos generados tienen diferentes estructuras tanto en la cabecera como en el cuerpo del mensaje transmitido, generando un riesgo de privacidad aumentando este desafío de interoperabilidad (Madaan, Ahad y Sastry, 2017).

### 1.3 Propuesta de solución

Después de analizar los trabajos previos, se identificó el reto de la interoperabilidad y la procedencia de los datos como parte de ella en las soluciones de IoT.

Pace *et al.* (2017) presentan su trabajo para apoyar la interoperabilidad de dos aplicaciones de asistencia médica Active and Assisted Living (AAL) mediante la integración de dos plataformas heterogéneas y no interoperables ya existentes como BodyCloud y universAAL. Este trabajo concluye con la propuesta del marco de trabajo INTER-IoT. Este proyecto tiene un horizonte hasta el 2020 y es un proyecto de la Comunidad Europea. Un caso de uso específico es INTERHealth que se despliega para validar el INTER-IoT propuesto. El proyecto estuvo compuesto por tres bloques de construcción principales: infraestructuras orientadas a capas (INTER-LAYER) para adaptar capas de pares heterogéneas (dispositivo a dispositivo, conexión de red a red, *middleware* a *middleware*, servicios de aplicaciones a servicios de aplicación, datos y semántica a datos y semántica (véase la figura 2).

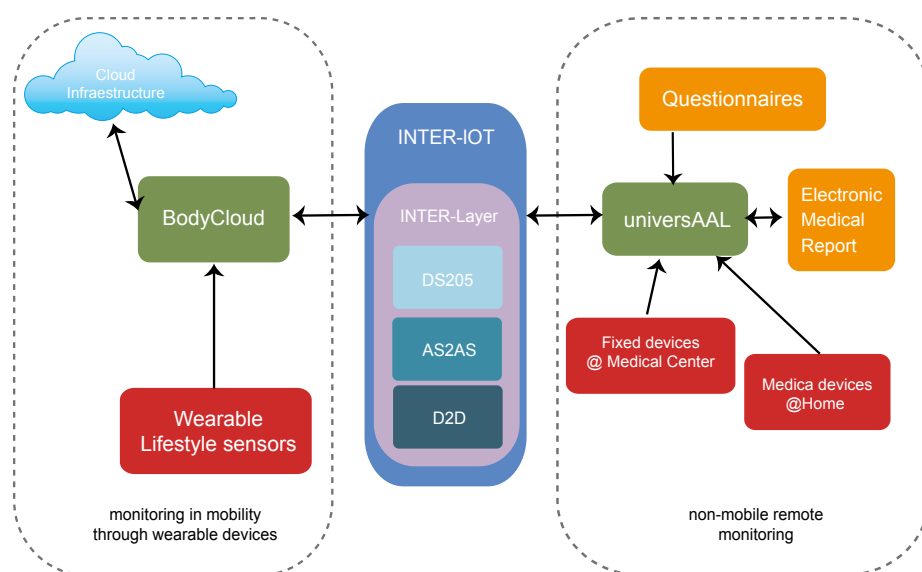


Figura 2. INTER-Health: integración BodyCloud y universAAL

Fuente: Datos tomados de Pace *et al.* (2017)

Di Martino, Esposito, Augusto Maisto, y Nacchia (2017) mencionan que existen varios actores que están en proceso de crear plataformas interoperables como M2M, AllJoyn, OIC y INTER-IoT, revisado líneas arriba. Presentan una propuesta de un marco de trabajo que permite una solución de interoperabilidad. Realizan una investigación sobre los diferentes IDL (*interface description language*) como los APIs WADL, RAML, Blueprint y OpenAPI para RESTful.

Di Martino *et al.*, (2017) presentan un *framework* como propuesta, que se aprecia en la figura 3, en el cual se muestra una arquitectura de alto nivel. Los componentes principales se destacan junto con los flujos de comunicación e interacción entre ellos. El *AgnosticEndPoint* es la interfaz que el usuario puede usar para realizar la solicitud. El *middleware* es el componente que es responsable de comprender si una entrada proporcionada por el usuario es útil para uno de los *endpoints* del proveedor registrado en el sistema; transformar la entrada proporcionada por el usuario en algo soportado por un “*endpoint*” del proveedor específico o múltiple; hacer la solicitud actual a *VendorEndPoint*; transformar la salida del *VendorEndPoint* en el formato estándar proporcionado por el *AgnosticEndPoint*.

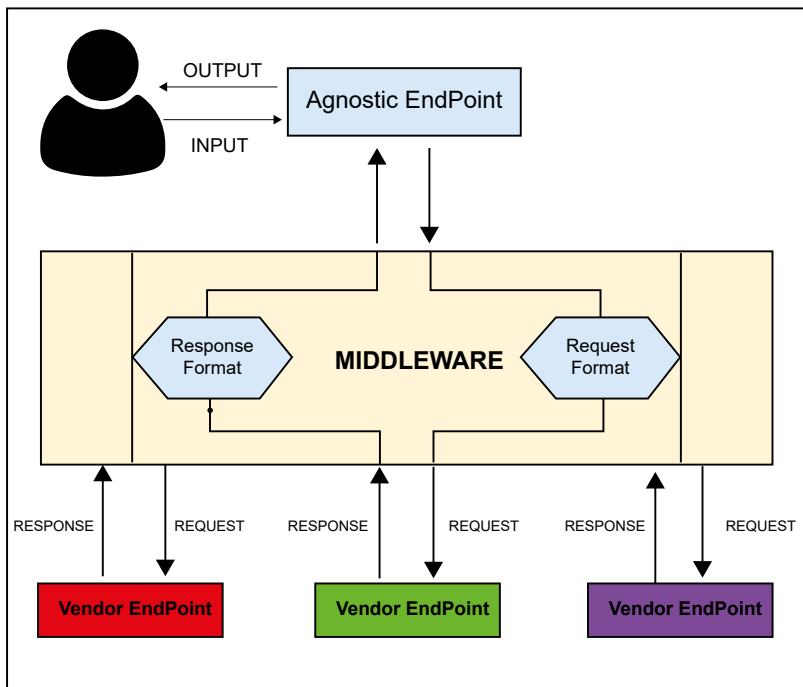


Figura 3. El flujo de proceso del *framework* presentado

Fuente: Di Martino *et al.* (2017)

Yachirema, Palau y Esteve (2017) proponen una arquitectura con dos capas: capa lógica de aplicación y capa de módulos comunes. La primera capa expone servicios; la segunda, cinco módulos básicos que habilitan la interoperabilidad de IoT para todos los servicios que se quieran supervisar, se la denomina Smart IoT Gateway y es la que aumenta la flexibilidad y escalabilidad del AAL-IoTSys propuesto. Esta arquitectura se desarrolló considerando que los dispositivos IoT, generalmente, son de recursos limitados o restringidos en términos de procesamiento, memoria y almacenamiento. La figura 4 muestra la arquitectura propuesta.

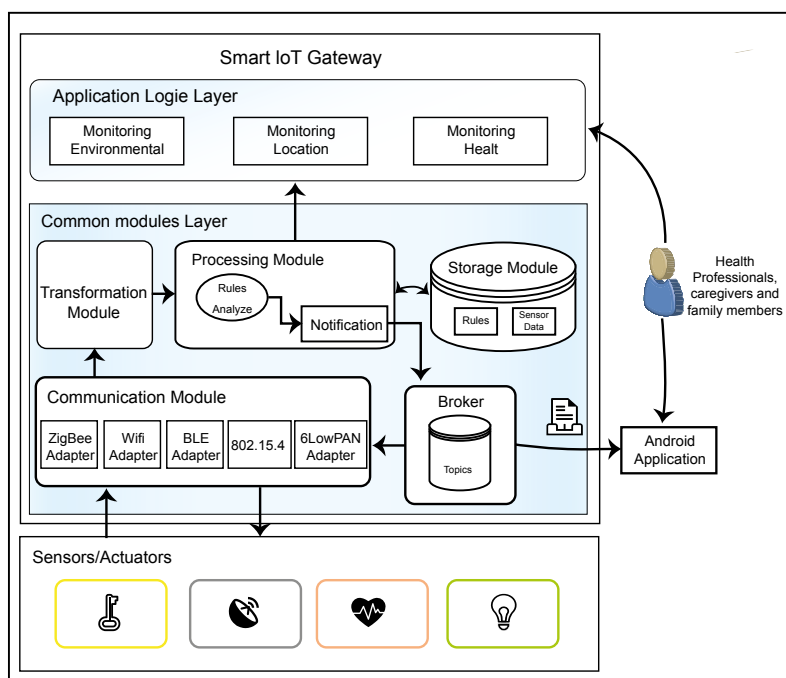


Figura 4. Descripción general de la arquitectura ALL-IoTSys

Fuente: Yachirema *et al.* (2017)

Andročec, Tomaš y Kišasondi (2017) plantean el uso de la web semántica para permitir la interoperabilidad en IoT utilizando JSON-LD. La idea básica de JSON-LD es crear una descripción de los datos en forma de un llamado contexto. También exploran algunas posibilidades de interoperabilidad y comunicación entre dispositivos IoT económicos.

Sun, Li y Memon (2017) formulan una REST API y una comparación entre la arquitectura monolítica y de microservicios. Utiliza REST API para la comunicación desde los dispositivos, incluyendo el registro del mismo en el ecosistema. La asignación de microservicios a dispositivos funciona dinámicamente al interactuar con el servicio central, proporciona un buen soporte para la (WoT) y tiene una mayor adaptabilidad, escalabilidad e interoperabilidad.

Lim, Majumdar y Nandy (2010) y Malik y Kim (2017) proponen y comparan arquitecturas basadas en servicios SOAP y REST API como propuesta de solución a la interoperabilidad.

Kum, Moon y Lim (2017) exponen una arquitectura novel para construir aplicaciones para *fog computing*, usando el procesamiento de datos en el borde en vez de hacerlo en la nube.

Petersen, Bindner, Poulsen y You (2017a) y Wendt, Faschang, Leber, Pollhammer y Deutsch (2013) hacen una demostración del desempeño de los diferentes formatos llegando a la conclusión de que el formato binario generado con Protobuff desarrollado por Google es el de mejor rendimiento y menor uso de memoria (véase la figura 5).

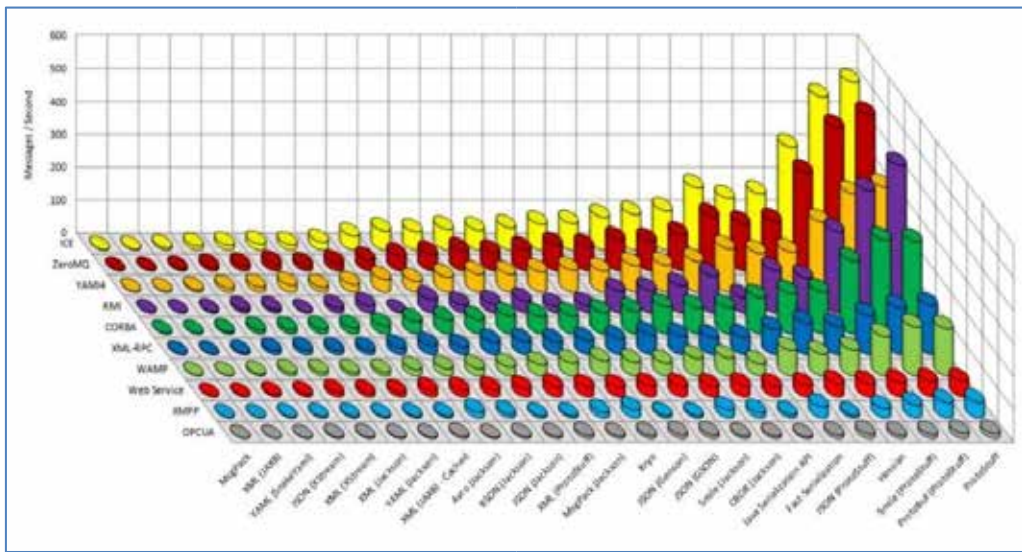


Figura 5. Prueba de rendimiento de request-response de formatos con diferentes protocolos

Fuente: Petersen *et al.* (2017a)

## 2. PROPUESTA

Hemos revisado varias propuestas de arquitectura de *software* para superar la interoperabilidad, así como de formatos de datos SOAP y REST API. Estas van desde microservicios y pasan por web semántica como JSON-LD, en ambos casos se requiere de capacidad de procesamiento y memoria que poseen los servidores, pero no dispositivos restringidos. Nuestro enfoque en este estudio es para dispositivos de bajo procesamiento y capacidad de memoria. La procedencia de los datos forma parte del reto de la interoperabilidad. Consideramos que existen ambientes hostiles de comunicación por lo que la trama de datos serializada debe ser muy compacta. La propuesta va dirigida a la capa al borde de un ecosistema de IoT (véase la figura 6).

Para nuestra propuesta tomamos como base el estudio comparativo de Petersen *et al.* (2017a) donde se demuestra un gran rendimiento utilizando el formato binario para la transferencia de datos desde los dispositivos.

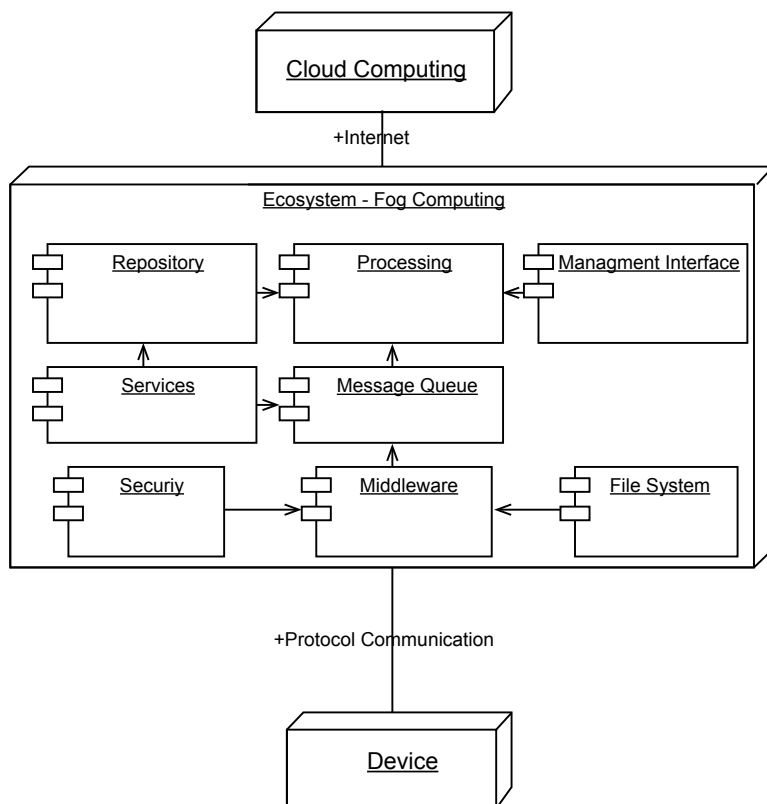


Figura 6. Propuesta de un patrón de arquitectura de *software* de IoT para superar la interoperabilidad en la capa del borde de un ecosistema IoT

Elaboración propia

## 2.1 Componentes que conforman la arquitectura

Describimos los componentes que conforman la arquitectura propuesta:

*Device*. Representa todos los dispositivos que se conectan al ecosistema a través del componente *middleware*. Aquí se debe implementar la serialización en formato binario. En general pueden ser sensores y actuadores.

- Sensor: detecta magnitudes físicas o químicas.
- Actuator: transforma la energía para causar un efecto en un elemento externo.

*Middleware.* Se encarga de recibir el formato binario. Este componente hace uso del componente *processing* que tiene mayor capacidad de procesamiento ya que se comporta como un servidor. Este componente también hace uso del componente *file system* como archivo de configuración del ecosistema.

*MessageQueue.* Se encarga de proporcionar alta disponibilidad y escalabilidad. Asegura que, ante la presencia de alta concurrencia, la trama de datos de los dispositivos se guarden siempre en la cola de mensajes para que de esa forma gestionar el encolamiento (Rostanski, Grochla y Seman, 2014).

*Security.* Este componente es un servicio consumido por el *middleware*. A cada *device* se le asigna un ID único mediante un certificado que luego será validado por este servicio.

*Processing.* Se encarga de procesar los formatos de datos binarios que son recibidos desde el componente *device*.

*Repository.* Es de persistencia donde se terminarán alojando las tramas de datos.

*Managment interface.* Interfaz de administración a la cual se puede acceder desde un terminal que puede ser un PC, una laptop o un celular. Mediante este administrador se registran los tipos de dispositivos, los dispositivos, *el middleware* y el ecosistema.

*File system.* Describe las propiedades del ecosistema y guarda los certificados de los ID de los *device* encriptados.

*Services.* Es la capa de servicios que el *fog computing* puede entregar para que otros dispositivos externos puedan recibir información (de sensores) o enviar una acción a ejecutar (actuadores).

*Fog computing.* Alberga componentes descritos. *Fog computing* tiene todas las ventajas del *cloud computing* (Paharia y Bhushan, 2018).

*Cloud computing.* Este componente, que está fuera del ecosistema de IoT, es el repositorio final y es donde se deben realizar el análisis y los cálculos de gran volumen de información. De tipo propietario de cientos de miles de servidores dedicados. De alta disponibilidad, permite que se puedan generar múltiples análisis de datos para diferentes fines (Mengistu, Alahmadi, Albuali, Alsenani y Che, 2018).

## 2.2 Descripción de interacción de los componentes

Describimos la interacción entre los componentes, lo primero que se tiene que hacer es identificar cada uno de los dispositivos, se observan los dispositivos D1, D2, D3, D4 y D5, cada uno de ellos tiene una identificación adicional que es el tipo de dispositivo (TD), vemos que los dispositivos D2 y D4 tienen el mismo tipo de dispositivo TD2, esto vendría a ser la categoría. En el dispositivo se deben definir dos bloques de datos, la cabecera de los datos donde se graba

el identificador del dispositivo que llamaremos ID y el identificador del tipo de dispositivo que llamaremos TID. En el cuerpo de los datos de la trama del dispositivo se registran los datos censados. Una vez enviada la trama de datos en formato binario, este llega al *middleware*. En el *middleware* lo primero que se hace es enviar la trama de datos al componente MessageQueue y luego se deserializa utilizando el componente *processing* y se valida que los datos de cabecera estén registrados en el componente *file system*. Previamente en el *file system* se ha registrado el tipo de dispositivo del cual pueden recibir la trama de datos y los dispositivos que le pueden enviar datos. En el componente *processing* se deben instalar los módulos de deserialización. Cuando se recibe un tipo de dispositivo el componente *processing* lo atenderá con el módulo que le corresponde, de esta forma si se registra un nuevo tipo de dispositivo se instalará solo el módulo de deserialización y se registrará en el *file system*. Esto permitirá realizar pruebas de concepto sin afectar a los otros módulos que ya están en operación y permitirá escalabilidad.

### 2.3 Diseño del *file system*

En nuestra base de datos NOSQL se crea la colección identificando el ecosistema al cual corresponde (véase la figura 9). Dentro de un ecosistema puede existir más de un *middleware*. Un *middleware* puede recibir tramas de varios tipos de dispositivos TD y puede haber varios dispositivos por tipo de dispositivo.

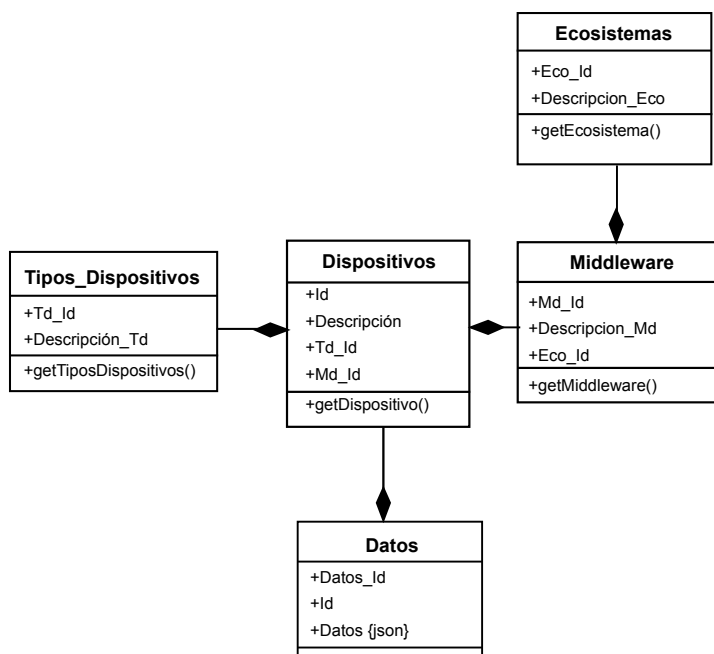


Figura 9. Diseño de datos del *file system*

Elaboración propia

Finalmente, si tuviéramos varios ecosistemas, cada uno tendría su identificación que luego podrían ser consolidados en la nube de tal manera que se aseguraría el escalamiento desde un *fog computing* hacia el *cloud computing*.

Por ejemplo:

ES1 / GW1 / TD1 / D1 / Trama-JSON

ES1 / GW1 / TD2 / D2 / Trama-JSON

ES1 / GW1 / TD2 / D4 / Trama-JSON

## 2.4 Escalando múltiples ecosistemas IoT mediante interoperabilidad

La figura 10 muestra la capacidad de interoperar entre los *device* con el ecosistema (*fog computing*), respetando el patrón propuesto tipificando los tipos de dispositivos, asignándole un comportamiento único en el proceso de serialización y deserialización. Esto facilita la integración de un nuevo dispositivo en el ecosistema IoT.

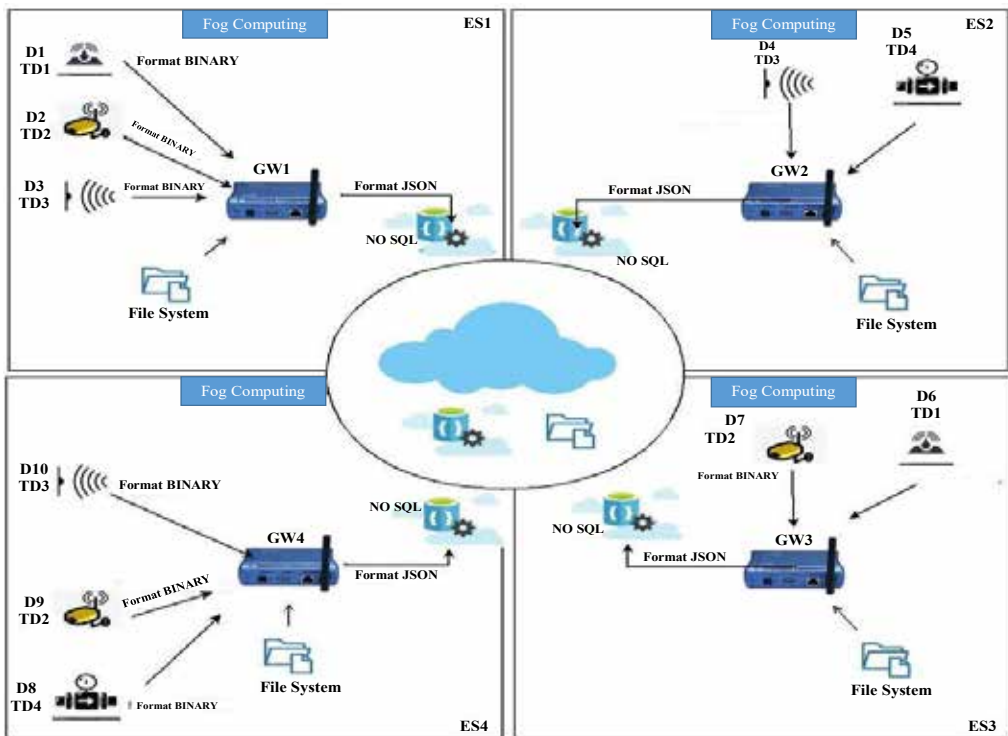


Figura 10. Integrando múltiples ecosistemas IoT

Elaboración propia



### 3. CONCLUSIONES

En este *paper* proponemos una nueva arquitectura de *software* para superar el reto de la interoperabilidad en soluciones IoT, iniciando con el diseño de un ecosistema (*fog computing*). Identificamos todos los componentes y sus interacciones. Diseñamos la estructura de datos que permitirá determinar la procedencia de los datos haciéndolos únicos dentro del ecosistema. Creamos ID únicos de los dispositivos a través de certificados encriptados para seguridad. Creamos la trama de formato binario para mejorar la velocidad de comunicación. Con este patrón logramos escalabilidad y alta disponibilidad integrándolos a través de un *cloud computing*.

### 4. TRABAJO FUTURO

Se ha implementado esta arquitectura en una aplicación en tiempo real como prueba de concepto. La solución implementa sensores para medir parámetros de medio ambiente tales como temperatura, humedad, sonido y monóxido, una minicomputadora Raspberri Pi 3 Model B (1.4 GH QuadCore) como *middleware*, una tarjeta Arduino, SQLite para el *file system*, Protocol Buffer para el *encode-decode* de los datos y el servicio de Firebase para hacer *cloud computing* de análisis de datos y almacenamiento.

Está pendiente desarrollar los otros componentes, como también el implementar dos o más ecosistemas con este patrón para demostrar la escalabilidad e identificar la procedencia de los datos mediante este diseño. El diseño de los datos de la prueba de concepto se puede ver en la figura 11.

Además, se debe realizar una prueba de estrés provocando concurrencia de los datos para la implementación del uso del componente MessageQueue (MQ) y asegurar la alta disponibilidad. Una primera implementación del uso de MQ se realizaría con RabbitMQ, sin embargo, luego estudiaremos la implementación de NATS y Kafka para manejo de MQ.

#### *Descripción de las columnas*

Los datos que se aprecian en la figura 11 se obtuvieron de un proceso de descarga que formatea en modo fila-columnas, pues los datos están guardados en formato JSON.

Detalle de los nombres de las columnas:

- *gwName*: nombre del *gateway*.
- *gwid*: identificador del *gateway*.
- *devName*: nombre del dispositivo.
- *devid*: identificador del dispositivo.

date	Oil temperature	Water concentration	Hydrogen	Coil temperature	Carbon monoxide	gwName	devName	gwId	devId
1/02/2019 5:10	27	6.4	35.1	67.8	273.6	SS.EE. Carapongo	CARP5000TR114-AUTOT-S	LisaGw	Lisa1_03
1/02/2019 5:34	27	6.4	35.1	67.9	273.6	SS.EE. Carapongo	CARP5000TR114-AUTOT-S	LisaGw	Lisa1_03
1/02/2019 5:46	26.8	6.4	35.1	67.8	273.6	SS.EE. Carapongo	CARP5000TR114-AUTOT-S	LisaGw	Lisa1_03
1/02/2019 5:58	26.8	6.4	35.1	67.8	273.6	SS.EE. Carapongo	CARP5000TR114-AUTOT-S	LisaGw	Lisa1_03
1/02/2019 6:10	26.7	6.4	35.1	67.8	273.6	SS.EE. Carapongo	CARP5000TR114-AUTOT-S	LisaGw	Lisa1_03
1/02/2019 6:22	26.7	6.4	35.1	67.8	280.8	SS.EE. Carapongo	CARP5000TR114-AUTOT-S	LisaGw	Lisa1_03
1/02/2019 6:34	26.7	6.4	35.1	67.8	280.8	SS.EE. Carapongo	CARP5000TR114-AUTOT-S	LisaGw	Lisa1_03
1/02/2019 6:46	26.7	6.4	35.1	67.8	280.8	SS.EE. Carapongo	CARP5000TR114-AUTOT-S	LisaGw	Lisa1_03
1/02/2019 6:58	26.7	6.4	35.1	67.8	280.8	SS.EE. Carapongo	CARP5000TR114-AUTOT-S	LisaGw	Lisa1_03
1/02/2019 7:10	26.7	6.4	35.1	67.8	280.8	SS.EE. Carapongo	CARP5000TR114-AUTOT-S	LisaGw	Lisa1_03
1/02/2019 7:22	26.7	6.4	35.1	67.8	280.8	SS.EE. Carapongo	CARP5000TR114-AUTOT-S	LisaGw	Lisa1_03
1/02/2019 7:34	26.7	6.4	35.1	67.8	280.8	SS.EE. Carapongo	CARP5000TR114-AUTOT-S	LisaGw	Lisa1_03
1/02/2019 7:46	26.5	6.4	35.1	67.8	280.8	SS.EE. Carapongo	CARP5000TR114-AUTOT-S	LisaGw	Lisa1_03
1/02/2019 7:58	26.5	6.4	35.1	67.8	280.8	SS.EE. Carapongo	CARP5000TR114-AUTOT-S	LisaGw	Lisa1_03

Figura 11. Ejemplo de datos de dispositivos sensados  
Elaboración propia

En nuestra prueba de concepto desarrollamos un dispositivo con múltiples sensores, esto permitió ahorrar en *hardware*.

## REFERENCIAS

- Alaya, M. B., Drira, K., y Gharbi, G. (2017). Semantic-aware Iot platforms. *2017 IEEE International Conference on AI & Mobile Services (AIMS)*. doi:10.1109/AIMS.2017.15
- Alkhalil, A., y Ramadan, R. A. (2017). IoT data provenance implementation challenges. doi:10.1016/j.procs.2017.05.436
- Andročec, D., Tomaš, B., y Kišasondi, T. (2017). Interoperability and lightweight security for simple IoT devices. *2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*. doi:10.23919/MIPRO.2017.7973621
- Bloebaum, T. H., y Johnsen, F. T. (2015). Exploring SOAP and REST communication on the Android platform. *MILCOM 2015 - 2015 IEEE Military Communications Conference*. doi:10.1109/MILCOM.2015.7357509
- Di Martino, B., Esposito, A., Augusto Maisto, S., y Nacchia, S. (2017). A semantic IoT framework to support RESTful devices' API interoperability. doi:10.1109/ICNSC.2017.8000071
- Dragoni, N., Giallorenzo, S., Lluch Lafuente, A., Mazzara, M., Montesi, F., Mustafin, R., y Safina, L. (2017). Microservices: yesterday, today, and tomorrow. doi:10.1007/978-3-319-67425-4\_12
- Elsts, A., Oikonomou, G., Fafoutis, X., y Piechocki, R. (2017). Internet of things for smart homes: lessons learned from the SPHERE case study. *Global Internet of Things Summit (GIoTS), 2017 Institute of Electrical and Electronics Engineers (IEEE)*. doi:10.1109/GIOTS.2017.8016226
- Gligorić, N., Dejanović, I., y Krčo, S. (2011). Performance evaluation of compact binary XML representation for constrained Devices. doi:10.1109/DCOSS.2011.5982183
- Guinard, D. D., y Trifa, V. M. (2016). *Building the web of things*. NY: Manning Publications.
- Jaimini, U. (2017). PhD Forum: Multimodal IoT and EMR based smart health application for asthma management in children. *2017 IEEE International Conference on Smart Computing (SMARTCOMP)*. doi:10.1109/SMARTCOMP.2017.7947025
- Jarwar, M. A., Ali, S., Kibria, G. M., Kibria, G., Kumar, S., y Chong, I. (2017). Exploiting interoperable microservices in web objects enabled internet of things. doi:10.1109/ICUFN.2017.7993746

- Kum, S. W., Moon, J., y Lim, T.-B. (2017). Design of fog computing based IoT application architecture. doi:10.1109/ICCE-Berlin.2017.8210598
- Lim, N., Majumdar, S., y Nandy, B. (2010). Providing interoperability for resource access using web services. doi:10.1109/CNSR.2010.23
- Lysogor, I., Voskov, L., y Efremov, S. (2018). Survey of data exchange formats for heterogeneous LPWAN-Satellite IoT networks. doi:10.1109/MWENT.2018.8337257
- Madaan, N., Ahad, M. A., y Sastry, S. M. (2017). Data integration in IoT ecosystem: Information linkage as a privacy threat. doi:10.1016/j.clsr.2017.06.007
- Malik, S., y Kim, D.-H. (2017). A comparison of RESTful vs. SOAP web services in actuator networks. *2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN)*, 753-755. doi:10.1109/ICUFN.2017.7993893
- Mengistu, T., Alahmadi, A., Albuali, A., Alsenani, Y., y Che, D. (2018). A “No data center” solution to cloud computing. doi:10.1109/CLOUD.2017.99
- Nakayama, M., Yamazaki, K., Tanaka, S., y Kasahara, H. (2014). Parallelization of Tree-to-TLV serialization. doi:10.1109/PCCC.2014.7017059
- Pace, P., Gravina, R., Aloï, G., Fortino, G., Fides-Valero, A., Ibañez-Sánchez, G., ... Yacchirema, D. (2017). IoT platforms interoperability for active and assisted living healthcare services support. doi:10.1109/GIOTS.2017.8016250
- Paharia, B., y Bhushan, K. (2018). Fog computing as a defensive approach against distributed denial of service (DDoS): a proposed architecture. *2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 1-7. doi:10.1109/ICCCNT.2018.8494060
- Petersen, B., Bindner, H., Poulsen, B., y You, S. (2017a). Smart grid communication comparison. Distributed control middleware and serialization comparison for the internet of things. *Proceedings of 7th IEEE International Conference on Innovative Smart Grid Technologies IEEE*. doi:10.1109/ISGTEurope.2017.8260268
- Petersen, B., Bindner, H., You, S., y Poulsen, B. (2017b). Smart grid serialization comparison. Comparison of serialization for distributed control in the context of the Internet of Things. *2017 Computing Conference*, 1339-1346. doi:10.1109/sai.2017.8252264
- Ray, P. P., Mukherjee, M., y Shu, L. (2017). Internet of things for disaster management: State-of-the-art and prospects. doi:10.1109/ICWS.2017.26
- Rostanski, M., Grochla, K., y Seman, A. (2014). Evaluation of highly available and fault-tolerant middleware clustered architectures using RabbitMQ. *2014 Federated Conference on Computer Science and Information Systems*, 879-884.

- Serrano, D., Stroulia, E., Lau, D., y Ng, T. (2017). Linked REST APIs: A middleware for semantic REST API integration. *2017 IEEE International Conference on Web Services (ICWS)*, 138-145. doi:10.1109/ICWS.2017.26
- Silverajan, B., Ocak, M., Jimenez, J., y Kolehmainen, A. (2016). Enhancing lightweight M2M operations for managing IoT gateways. *9th IEEE International Conference on Internet of Things (iThings 2016)* (pp. 187-192). IEEE. doi:10.1109/iThings-GreenCom-CPSCoM-SmartData.2016.55
- Sun, L., Li, Y., y Memon, R. A. (2017). An open IoT framework based on microservices architecture. *China Communications*, 14, 154-162. doi:10.1109/CC.2017.7868163
- Talavera, J. M., Tobón, L. E., Gómez, J. A., Alejandro, M. A., Aranda, J. M., Parra, D. T., ... Garreta, L. E. (2017). Review of IoT applications in agro-industrial and environmental fields. *Computers and Electronics in Agriculture*, 142, 283-297. doi:10.1016/j.compag.2017.09.015
- Wendt, A., Faschang, M., Leber, T., Pollhammer, K., y Deutsch, T. (2013). Software architecture for a smart grids test facility. *IECON 2013 - 39th Annual Conference of the IEEE Industrial Electronics Society*, 7062-7067. doi:10.1109/IECON.2013.6700304
- Yacchirema, D. C., Palau, C. E., y Esteve, M. (2017). Enable IoT interoperability in ambient assisted living: active and healthy aging scenarios. *2017 14th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, 53-58. doi:10.1109/CCNC.2017.7983081
- Yigitoglu, E., Liu, L., Looper, M., y Pu, C. (2017). Distributed orchestration in large-scale IoT systems. *2017 IEEE International Congress on Internet of Things (ICIOT)*, 58-65. doi:10.1109/IEEE.ICIOT.2017.16



# Generación de reglas de asociación para productos de *retail* utilizando el algoritmo FP-Growth paralelo

Renato Pérez-Gómez  
renato08lasalle13@gmail.com / Universidad de Lima, Perú

Recepción: 19-6-2019 / Aceptación: 8-8-2019

RESUMEN. Las organizaciones minoristas actuales tienen varias sucursales conectadas bajo el mismo sistema de gestión distribuido. Estos sistemas almacenan y registran la información de todas las transacciones dadas en las tiendas. Paralelamente, con el rápido crecimiento e implementación de las tecnologías de la información e Internet, la cantidad de datos generados en cada transacción o venta realizada es sustancial. Las técnicas de minería de datos tienen como objetivo identificar patrones y tendencias en una gran recopilación de datos. Su uso tiene un atractivo para los minoristas, ya que quieren convertir la gran cantidad de datos que tienen en información y conocimiento útiles. Una aplicación de minería de datos que atrae a los minoristas es el descubrimiento de reglas de asociación. El descubrimiento de estas reglas es la base de muchas decisiones comerciales, como el diseño de la canasta de productos, la elección de la estrategia de promoción y la combinación de productos. La intención de la investigación es la aplicación de las técnicas y metodologías de aprendizaje de asociación de reglas para la realidad de un comercio minorista con sucursales ubicadas en Lima. El conjunto de datos utilizado en esta investigación corresponderá a las transacciones realizadas con el tiempo para productos de consumo masivo.

PALABRAS CLAVE: minería de datos, análisis de la cesta de la compra, análisis de reglas de asociación, conjuntos de elementos frecuentes, FP-Growth, computación paralela

## Generation of Association Rules for Retail Products Using the Parallel FP-Growth Algorithm

**ABSTRACT.** Retail organizations today have several branches connected under the same distributed management system. These systems store and record the information of all transactions taken place in stores. In parallel, with the rapid growth and implementation of information technologies and the Internet, the amount of data generated in each transaction or sale is substantial. The data mining techniques are aimed at identifying patterns and trends in a large data collection. The use of these techniques is attractive for retailers, as they want to convert the large amount of data they have into useful information and knowledge. A data mining application that attracts retailers is the discovery of association rules. The discovery of these rules is the basis of many commercial decisions, such as the design of the product basket, the choice of the promotion strategy and the combination of products. This research aims to apply learning techniques and methodologies of association rules for a retailer with branches located in Lima. The data set used in this research will match the transactions made over time for mass consumption products.

**KEYWORDS:** data mining, market basket analysis, association rule analysis, frequent itemsets, FP-growth, parallel computing



## 1. INTRODUCCIÓN

Con la tendencia, en los últimos años, de registrar y almacenar todo dato posible relacionado a los procesos operativos de la organización, la cantidad de *data* generada es abundante. Respecto a la industria del *retail*, gran parte de la *data* generada proviene de los miles de transacciones efectuadas en cada una de las sucursales. Los grandes representantes del sector a nivel mundial han identificado el gran potencial en minar la información oculta dentro de estos conjuntos de datos. Sin embargo, la aplicación de las técnicas de minería de datos no es el estándar en la mayoría de las organizaciones. La mayor aplicación de la minería de datos en el sector *retail* es la minería de reglas de asociación. Esta práctica supone un atractivo para las organizaciones debido a que tiene la intención de convertir la gran cantidad de datos que poseen en información útil y conocimiento. El descubrimiento de estas reglas es la base de muchas decisiones comerciales, como el diseño de la cesta de productos, la elección de la estrategia de promoción y la combinación de productos. Para la minería de reglas de asociación normalmente se hace uso del algoritmo Apriori. No obstante, en el desarrollo del algoritmo no estuvo pensado para manejar grandes volúmenes de datos por lo que en la actualidad es costoso, tanto computacional como en memoria. Existen otras alternativas óptimas respecto al manejo de memoria y que requieren de menos capacidad computacional; el algoritmo FP-Growth demuestra tener mejores resultados en rendimiento. Este algoritmo busca resolver las desventajas que posee el algoritmo clásico. Ya que este consiste en dos pasos para la generación de las reglas de asociación, es menos costoso en tiempo y en recursos. Si a este algoritmo se le suma la capacidad de procesar información en paralelo, el tiempo necesario para la ejecución del algoritmo será reducido drásticamente. Para lograr esto, en esta investigación, se hizo uso del *framework* Apache Spark y de su implementación del algoritmo FP-Growth.

## 2. ESTADO DEL ARTE

Los algoritmos existentes para la minería de reglas de asociación funcionan en datos estáticos, (Kaur y Kang, 2016). Estos encuentran las mejores reglas de asociación en base a métricas como soporte, confianza, elevación, etc. Sin embargo, para la próxima vez que se realiza la extracción de datos, estos algoritmos no capturan automáticamente los cambios en los datos. Por ello se usa algún otro algoritmo de comparación para rastrear el cambio en los datos. Este último, se utiliza para comprender la dinámica del proceso de generación de datos mediante la examinación de los cambios que se han producido en los patrones descubiertos. Se concluye que la minería periódica es un nuevo enfoque en la minería de datos que adquiere importancia. Este campo está evolucionando debido a las necesidades en diferentes aplicaciones y limitaciones de la minería de datos. Esto aumentaría el poder de las técnicas existentes de extracción de datos.

Según Di Fatta (2019), podría decirse que algunos factores limitantes han impedido una adopción más generalizada de ARM (*association rule mining*). En primer lugar, el descubrimiento

de patrones interesantes a partir de datos requiere algoritmos combinatorios complejos y, a menudo, se benefician de la computación de alto rendimiento. Para problemas del mundo real, el espacio de búsqueda puede ser prohibitivamente grande. En segundo lugar, no es raro que incluso para un conjunto de datos de entrada relativamente pequeño, el conjunto de patrones descubiertos sea muy grande, incluso más grande que el conjunto de datos de entrada. En este caso, ARM es solo un primer paso en flujos de trabajo de datos más complejos que incluyen otras técnicas de extracción de datos. Por estas razones, un enfoque multidisciplinario que combina la experiencia tanto de la informática como del dominio de la aplicación específica es a menudo crítico.

En la investigación realizada por Griva, Bardaki, Pramatarí y Papanikolaou (2018) se propone un enfoque de análisis de negocios que extrae segmentos de visitas de clientes a partir de datos de venta. Se caracteriza por visita al cliente según las categorías de productos comprados en la canasta e identifica la intención de compra o la misión detrás de la visita. Además, se sugiere un enfoque de selección de funciones semisupervisadas que utiliza la taxonomía del producto como entrada y sugiere categorías personalizadas como salida. Este enfoque se utiliza para equilibrar el árbol de taxonomía del producto que tiene un efecto significativo en los resultados de la extracción de datos. En general, el resultado de esta investigación permite observar el comportamiento de compra de los clientes desde un punto de vista alternativo, no como sus necesidades e intenciones generales de compra, sino que se centran en sus requisitos y motivos por cada viaje o visita de compras.

Referente a Bhandari, Gupta y Das (2015), identifican la principal limitación de Apriori como el costo de tiempo para mantener un gran número de conjuntos candidatos con conjuntos de elementos frecuentes, soporte mínimo o conjuntos de elementos grandes. Por lo que proponen modificar el algoritmo para generar un FP-Tree en lugar de iterar el *dataset* una gran cantidad de veces. Con el algoritmo de Apriori se escaneó la base de datos dos veces, pero evidenciaron una mejora al usar el algoritmo paralelo y el concepto de partición. Presentan una fórmula matemática para seleccionar el clúster, ya que hay muchos clústeres. El resultado, un híbrido de Apriori con generación de FP-Tree. Concluyen, así, que el espacio de memoria se reduce drásticamente cuando se realiza una gran cantidad de transacciones desde los *data-warehouse* y al reducir el tiempo consumido en el escaneo de transacciones y también al reducir el número de transacciones a escanear.

En Galarreta Vásquez (2016) se identifican reglas de asociación entre categorías de productos de electrodomésticos de un *retail*. Para la prueba utilizaron *data* histórica de diez años para tres tiendas. El modelo se construyó con el *software* RapidMiner conectado a una BD en SQL Server. Eligió el algoritmo FP-Growth al ser más eficiente que Apriori e hizo uso de la totalidad de registros disponibles en la BD. Los productos presentaban la siguiente jerarquía: familia, subfamilia, marca y oferta. La *data* fue filtrada por oferta para dividir los productos en campañas promocionales, ya que podían alterar el resultado. La BD fue transformada para

generar una matriz, la cual, luego, fue procesada por el algoritmo. La matriz almacenó la información cruzada de todo cliente que compró en el local y todas las subfamilias de productos. Si un cliente compró, por lo menos una vez un producto de la subfamilia, entonces fue registrado como TRUE, sea contrario como FALSE. Se concluyó que los productos con bajo precio incentivan la compra de productos con mayor precio.

Respecto a Xue, Wang, Liu y Li (2010), diseñan un modelo de minería de reglas de asociación para una venta de libros *online*. Para esto tienen 623 registros en la BD, los cuales fueron separados en 4 tablas. Definieron los parámetros soporte y confianza con 2 % y 60 %, respectivamente. Para la minería crearon una librería basada en ASP.NET. Luego, con un experimento analizaron los hábitos de compra de los clientes al encontrar asociaciones entre los artículos de pedido en la base de datos. La información obtenida del resultado del ejercicio concluyó que se lleva un aumento de las ventas al ayudar a los *retailers* a realizar una comercialización selectiva y planificar el espacio de la librería.

En Chunhua y Dongjun (2008) se presentan dos modelos de minería capaces de adaptarse a dos tipos de organización. Los autores plantean que existen dos tipos de organizaciones, aunque todas adoptan el enfoque de minería distribuida y el mismo patrón de toma de decisiones. En general, la carga de comunicación y la frecuencia de comunicación son los dos factores que afectan la eficiencia de la minería de datos. Se puede analizar un sistema de gestión distribuido desde tres aspectos: 1) cantidad de la tienda de sucursales; 2) el ancho de banda de comunicación; 3) entre las características del conjunto de datos local, las características incluyen si es denso o escaso, las cantidades de datos de transacciones, etc. Se definen dos tipos de organizaciones de *retail*. El primer tipo corresponde a las organizaciones que tienen varias sucursales y el bando de ancho de comunicación es largo (gran escala), mientras que la *data* de transacción de cada sucursal es pequeña. El segundo tipo, tienen pocas sucursales y comunicación pequeña, mientras que la *data* de transacción es grande (pequeña escala).

### 3. ANTECEDENTES

La minería de *itemsets* frecuentes dentro del *dataset* es el punto clave de la generación de reglas de asociación. Lo que se pretende con esta técnica es extraer los *itemsets* más frecuentes y grandes dentro de una gran lista de transacciones que contienen varios elementos cada uno. A continuación, una breve descripción de dos algoritmos para la minería de reglas de asociación.

#### 3.1 Algoritmos de minería de asociación

La primera propuesta de solución presentada se dio en Agrawal, Imienlinski y Swami (1993), en donde se presenta el algoritmo Apriori. Durante veinte años, este algoritmo se presentó como la solución clásica en el área. Sin embargo, muchas otras soluciones han sido propuestas

para obtención de las reglas de asociación para base de datos transaccionales. Las dos grandes desventajas que se pueden identificar de Apriori son el tiempo de ejecución que crece exponencialmente con la cantidad de registros y los requisitos computacionales para la generación de estas reglas. Ante estos puntos, se presenta el algoritmo FP-Growth como la alternativa ideal (Han, Pei y Yin, 2000).

### 3.2 FP-Growth

Es una mejora de Apriori diseñada para eliminar algunos de los cuellos de botella en este (Han *et al.*, 2000). El algoritmo se diseñó teniendo en cuenta los beneficios de mapReduce, por lo que funciona bien con cualquier sistema distribuido centrado en mapReduce. FP-Growth simplifica todos los problemas presentes en Apriori mediante el uso de una estructura llamada FP-Tree. En un árbol FP, cada nodo representa un elemento y su recuento actual, y cada rama representa una asociación diferente.

El árbol FP se construye leyendo el conjunto de datos una transacción a la vez y asignando cada transacción a una ruta en el árbol de FP (Han, Kamber y Pei, 2011). Como diferentes transacciones pueden tener varios elementos en común, sus rutas pueden superponerse (Tan, Steinbach y Kumar, 2005). Cuanto más se superponen las rutas, más compresión se puede lograr utilizando la estructura del árbol de FP. Si el tamaño del árbol de FP es lo suficientemente pequeño como para caber en la memoria principal, esto permitirá extraer conjuntos de elementos frecuentes directamente de la estructura en la memoria en lugar de realizar pases repetidos sobre los datos almacenados en el disco.

Para minar el árbol FP se empieza desde cada patrón de longitud 1 (patrón sufijo inicial), se construye su base de patrón condicional (consiste en el set de rutas de prefijo en el árbol que concurre con el patrón sufijo), luego se construye su árbol FP (condicional), y se realiza la minería de forma recursiva en el árbol (Han *et al.*, 2011). El crecimiento del patrón se logra mediante la concatenación del patrón sufijo con los patrones frecuentes generados a partir de un árbol FP condicional.

La mayor ventaja que se encuentra en FP-Growth es el hecho de que el algoritmo solo necesita leer el archivo dos veces, como se explica a continuación, quien lo lee una vez por cada iteración. Otra gran ventaja es que elimina la necesidad de calcular los pares que se van a contar, lo cual es muy pesado en el procesamiento, ya que utiliza el FP-Tree (Fournier-Viger *et al.*, 2017). Esto lo hace  $O(n)$  que es mucho más rápido que Apriori. El algoritmo FP-Growth almacena en memoria una versión compacta de la base de datos.

## 4. METODOLOGÍA

### 4.1 Exploración de la *data*

Antes de entrar a detallar el modelo y su funcionamiento, es importante dedicar esfuerzos a entender la *data* con la que se está trabajando. El análisis exploratorio de datos se refiere al proceso crítico de realizar investigaciones iniciales sobre datos para descubrir patrones, detectar anomalías, probar hipótesis y verificar suposiciones con la ayuda de estadísticas de resumen y representaciones gráficas. Es una buena práctica comprender los datos primero y tratar de recopilar tantos conocimientos a partir de ellos.

Luego de haber realizado la limpieza necesaria al *dataset* para que este se encuentre listo y sea alimentado al modelo, se realiza un análisis exploratorio a la *data*. De esta manera, la información que se obtenga permitirá generar visión de negocio. Es decir, conocimiento sobre cuál es el comportamiento entre las entidades del caso de estudio. Para este caso en específico, sería la relación entre productos y departamentos o categorías, los diez productos con más órdenes, las horas o días de la semana con más órdenes, etcétera.

#### 4.1.1 Distribución de productos por categoría

Esta observación permite entender cuáles son los productos que más espacio ocupan en estantes y a qué categoría pertenecen (véase la figura 1). El espacio en los estantes influye en la efectividad de ventas y *marketing*. Evidentemente, existe una diferencia en la visibilidad y atención recibida entre los estantes inferiores y los estantes superiores, por lo que se puede observar una diferencia significativa entre las ventas de los productos y marcas colocadas en dos estantes diferentes. Esta distribución es fundamental para que los *retailers* puedan establecer sus estrategias de negocio y *marketing*.



Figura 1. Distribución de productos por categoría  
Elaboración propia

#### 4.1.2 Productos más vendidos

Es importante también tener en cuenta cuáles son los productos que más ventas u órdenes registran ya que son estos los puntos clave para las estrategias de negocio que elaboren los *retailers*. En la figura 2 se pueden observar los diez productos más vendidos. Esto, junto con el análisis anterior, permite saber más si la distribución de ventas de la categoría depende de uno o un grupo selecto de productos o del conjunto total.

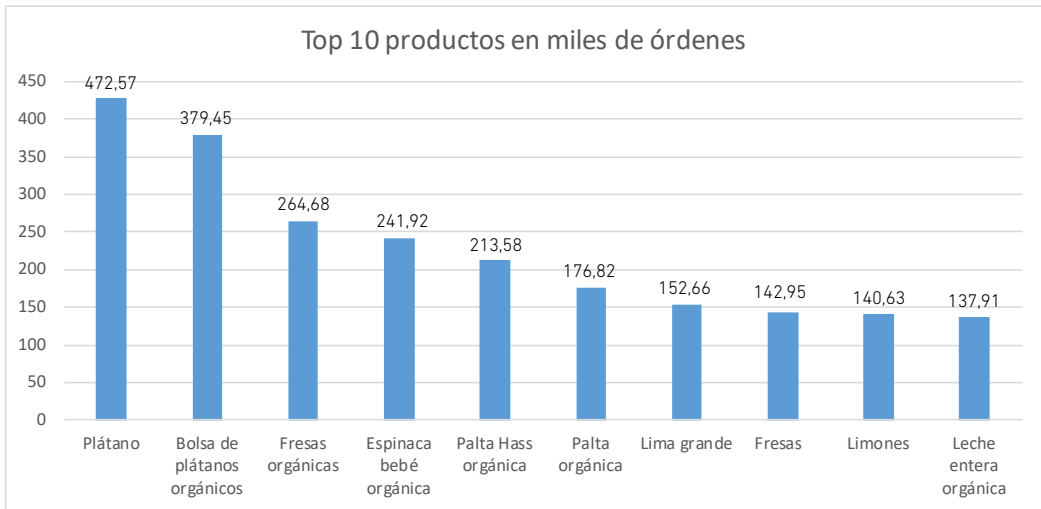


Figura 2. Top 10 productos en miles de órdenes  
Elaboración propia

#### 4.1.3 Número de productos por transacción u órdenes

De igual manera que las visualizaciones anteriores, otro aspecto importante a tomar en cuenta es la cantidad de productos que se registran por transacción u orden. Esto quiere decir, cuántos productos compra cada cliente. De esta forma, los *retailers* podrán crear ofertas personalizadas que logren atraer más a un cliente según esta métrica.

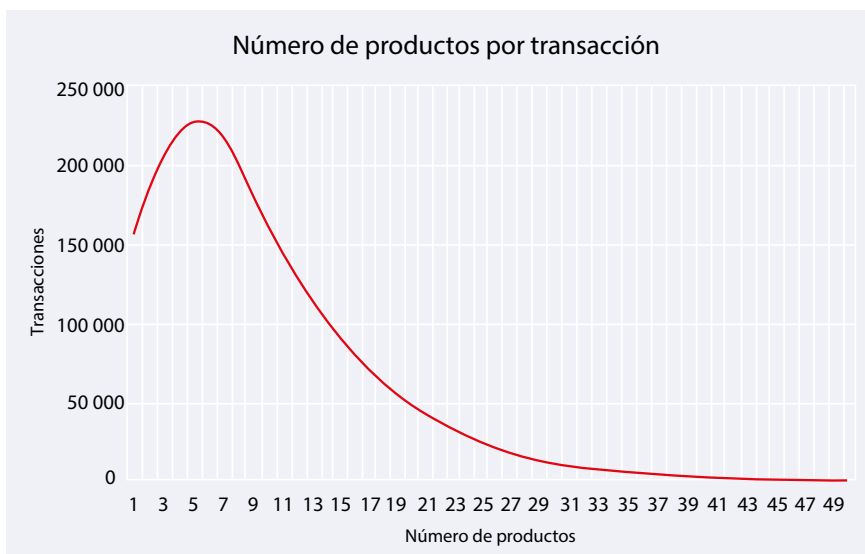


Figura 3. Número de productos por transacción  
Elaboración propia

#### 4.1.4 Cantidad de órdenes en el tiempo

Por último, esta visualización permite conocer las frecuencias de compra de los clientes. Qué días de la semana y a qué hora de dicho día se registran las mayores cantidades de transacciones. Con estos datos es posible determinar la probabilidad de que un cliente existente con un historial de compras adquiera un producto en un momento determinado. Esta información permite no solo dirigirse a los clientes que tienen más probabilidades de comprar algo, sino también adaptar la oferta a lo que es más probable que les atraiga.

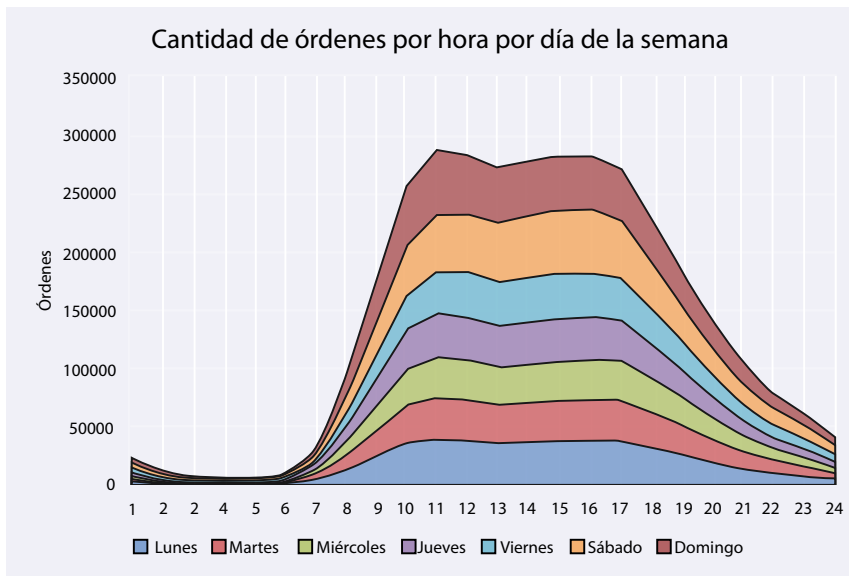


Figura 4. Cantidad de órdenes por hora por día de la semana

Elaboración propia

## 4.2 Modelo y experimentación

Se tomó la decisión de desarrollar la propuesta de solución en el lenguaje Python. Este es el lenguaje más utilizado en el área de Data Science (por lo que tiene una excelente comunidad y extenso conjunto de librerías), es liviano y eficiente en la ejecución de código, pero también es multifuncional. De igual forma, se utilizó el *framework* de computación en clúster y de código abierto Apache Spark, este provee una interfaz para la programación con paralelización de la *data* implícita y tolerancia a fallos.

La legibilidad y la simplicidad inherentes de Python hacen que sea relativamente fácil de recopilar y la cantidad de bibliotecas analíticas dedicadas disponibles en la actualidad significa que se puedan encontrar paquetes ya adaptados a las necesidades del problema que se pueden descargar libremente. Esto se debe a que Python es un lenguaje de programación



multiparadigma: una especie de navaja suiza para el mundo de la codificación. Es compatible con la programación orientada a objetos, la programación estructurada y los patrones de programación funcional, entre otros.

Apache Spark es un potente motor de código abierto que ofrece procesamiento de flujo en tiempo real, procesamiento interactivo, procesamiento de gráficos, procesamiento en memoria y procesamiento por lotes con una velocidad rápida, facilidad de uso e interfaz estándar. Dentro de los componentes existentes en el ecosistema de Spark, en esta propuesta de solución se utilizarán los módulos MLlib y SQL. Esta interfaz contiene implementaciones de algoritmos y modelos especializados de *machine learning*. En este módulo se encuentra la implementación del algoritmo FP-Growth.

Respecto al modelo planteado en este trabajo para la generación de reglas de asociación utilizando el algoritmo FP-Growth, lo primero que se realizó fue identificar los módulos necesarios para la solución. Estos son, un módulo para el preprocesamiento y exploración de la *data* y otro módulo para el algoritmo de generación de ítems frecuentes y generación de reglas de asociación. Lo que en el área de *machine learning* se conoce como *pipelines*. Esto se refiere a la estructura secuencial de pasos a seguir para la implementación del modelo, siendo la salida de cada paso la entrada del siguiente.

El flujo de trabajo representativo de la propuesta de solución se puede observar en la figura 5. A partir de lo explicado en esta sección, se pueden abstraer cinco pasos críticos que contienen los procedimientos necesarios a llevar a cabo desde la lectura del *dataset* (*input*) hasta la generación de las reglas de asociación (*output*).

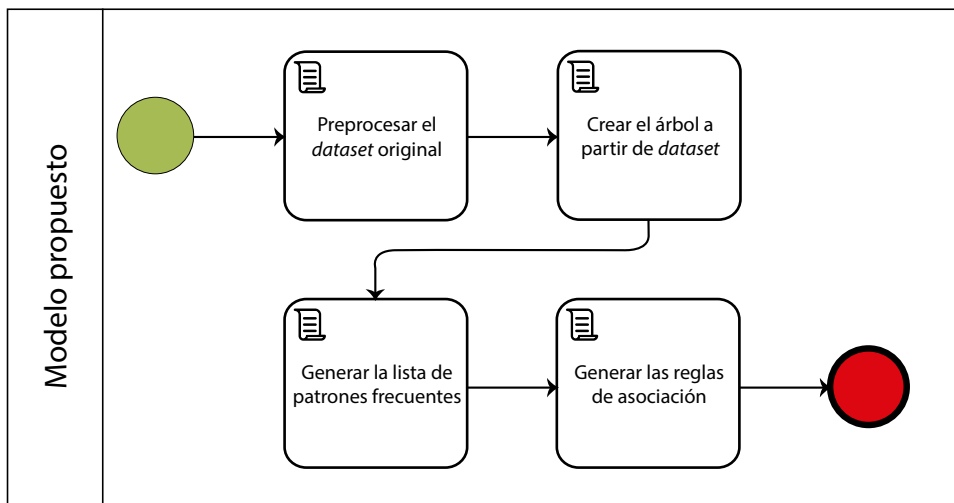


Figura 5. Flujo del modelo propuesto

Elaboración propia

Antes de describir en qué consisten los pasos mencionados, es necesario entender la estructura del *dataset* que se posee, porque esta estructura define el enfoque a tomar para las acciones respectivas al preprocesamiento, a cómo realizar un análisis exploratorio y cómo leer esta *data*, y la estructura que este conjunto de datos debe tener para poder ser alimentado al modelo. El *dataset* en posesión es una fracción en dimensión de tiempo de los datos almacenados en una base de datos (BD) relacional para un *retailer* del sector autoservicio. El nombre del *retailer* permanecerá confidencial por mutuo acuerdo. Sin embargo, es posible detallar que esta fracción pertenece a registros correspondientes a la temporada verano. Este dato será clave en el apartado de resultados, ya que se da contexto al *output* del modelo. Al ser un reflejo de la BD, se tiene un archivo de tipo *csv* para cada entidad, en total cuatro: pasillos, categorías, órdenes y productos. Estas se relacionan a partir de identificadores (Id) únicos para cada registro de cada tabla, como una BD transaccional.

Para el módulo de preprocesamiento se eliminaron todos los registros NA, como primer paso. Después se procedió a determinar las variables o campos factores, que en este caso vendrían a ser ProductID y OrderId. Estas variables permiten la agrupación de los ítems por transacción. Esto quiere decir, si  $n$  número de ítems pertenecen a la misma transacción  $t$ , esto se traducirá a una lista de transacciones donde  $t$  contendrá los  $n$  ítems. Finalmente, ya que el modelo necesita una lista de tuplas compuestas del índice de la transacción y de la lista de ítems en esa transacción, se selecciona la lista de ítems producto de la agrupación realizada en el paso anterior. Asimismo, para la exploración de la *data* se utilizó el componente SQL de Spark. Este permite realizar consultas SQL a los *dataframes* de Spark, algo que facilita el tratamiento de la *data* ya que esta posee una estructura relacional. No se detallarán los resultados de la exploración de la *data* dado que eso fue realizado en el apartado 4.1.

Con respecto al módulo para la generación de los ítems frecuentes y las reglas de asociación, la implementación del algoritmo en Spark requiere de cuatro parámetros, los cuales son la columna en el *dataset* que contiene a la lista de ítems por transacción, el soporte mínimo, la confianza mínima y el número de particiones a realizar el paralelismo del procesamiento. La columna para evaluar es simplemente el nombre de la columna en el *dataframe* Spark que contiene la lista de ítems. El soporte y la confianza mínimos son valores que van en el rango de 0 a 1, y se expresan en porcentaje en el análisis. Los valores para estos dos parámetros dependen de las necesidades de la empresa y es ella la encargada de realizar el estudio correspondiente para establecer estos valores. Debido a que el análisis de las necesidades de la empresa es personalizado e implica muchos aspectos a considerar, este escapa del alcance de esta investigación. El número de particiones es necesario cuando se desea ejecutar el modelo creado en un clúster de computadoras. Puesto que en este trabajo todo procesamiento es ejecutado desde un solo ordenador, este parámetro no impacta en el desempeño real del modelo. Posteriormente, ya definidos los parámetros para el modelo, se pasa a realizar el ajuste de este con el conjunto de datos especificado. El resultado será el modelo adaptado a los valores dados a partir de la *data* alimentada. Una vez realizado esto, se pueden obtener los ítems frecuentes y reglas de asociación gracias a las funciones ya definidas en el *framework* de Apache Spark.

La abstracción de la lógica detrás de la generación de ítems frecuentes y reglas de asociación, se definen en los algoritmos *find\_frequent\_patterns* y *generate\_association\_rules*. La última es la que genera las reglas de asociación a partir de los patrones frecuentes encontrados. Respecto al algoritmo *find\_frequent\_patterns*, se instancia un objeto de clase árbol con los parámetros transacciones y umbral de soporte mínimo para filtrar los patrones frecuentes. Este objeto hará llamado a la función *mine-tree*, esta función actuará dependiendo de la condicional de si el árbol contiene un solo camino o rama que nace desde el nodo o contiene múltiples caminos o ramas. Si el árbol tiene un solo camino se generará la lista de patrones frecuentes. Si el árbol tiene muchos caminos se minarán estos subárboles y se tendrá una sublista de patrones por subárbol, estas sublistas luego serán insertadas en el diccionario principal de patrones, el cual tendrá aquellos patrones por subárbol que superen el mínimo de soporte establecido. Con respecto al manejo de la clase árbol, este se generará y trabajará de manera común, con recursividad. Respecto al algoritmo *generate-association-rules*, tomará el diccionario que generó *find-frequent-patterns* y lo iterará para generar un diccionario que tenga como clave al antecedente y como valores al consecuente y a la confianza. La confianza representa la probabilidad de que el consecuente se presente cuando exista el antecedente.

---

ALGORITHM 1: *find-frequent-patterns* (transacciones, soporte mínimo)

---

```

Tree <- Instanciar un objeto árbol con parámetros el dataset y el soporte mínimo deseado
IF tree tiene un solo camino:
    Generar diccionario de patrones frecuentes
ELSE:
    Generar sublista de patrones frecuentes por cada subárbol
Insertar sublista de patrones frecuentes en el diccionario principal de patrones frecuentes
Output <- diccionario de patrones frecuentes
End

```

---

ALGORITHM 2: *generate-association-rules* (patrones, confianza mínima)

---

```

FOR itemset IN patrones.clave:
    soporte_superior <- patrones[itemset]
FOR index IN itemnset.longitud:

```

```

FOR antecedente IN itemset[i]:
    IF antecedente se encuentra en patterns:
        soporte_inferior <- patrones[antecedente]
        confianza <- soporte_superior / soporte_inferior
        IF confianza >= confianza_mínima:
            Reglas[antecedente] = (consecuente, confianza)
Output <- Reglas
End
    
```

---

ALGORITHM 3: FP-Growth

---

```

Patrones <- find_frequent_patterns(dataset, soporte_min)
Reglas <- generate_association_rules(patrones, confianza_min)
Output <- Reglas
End
    
```

## 5. RESULTADOS Y DISCUSIÓN

El modelo se inicializó con los valores 0,01 para soporte mínimo y 0,1 para confianza. Los ítems frecuentes encontrados que sobrepasen el umbral de soporte están conformados por dos ítems. Esto quiere decir que no existen relaciones entre 3 ítems que pasen el 1 % del total de transacciones. En la figura 6 se puede apreciar los 16 primeros ítems frecuentes según su frecuencia. La relación más fuerte existente entre fresas orgánicas y bolsa de plátanos orgánicos. Existen un total de 131 209 transacciones, por lo que la relación descrita aparece el 2,34 % de veces.

Referente a las reglas de asociación, la relación más fuerte es entre palta Hass orgánica (antecedente) y bolsa de plátanos orgánicos (consecuente). Esta relación se refiere a la probabilidad de que ocurra el consecuente a partir del antecedente. Es decir, hay un 33,18 % de que la bolsa de plátanos orgánicos sea comprado cuando se compra palta Hass orgánica. A diferencia de los ítems frecuentes, este examina la relación pirobalística entre dos o más ítems, mientras que los ítems frecuentes es una descripción frecuencial de las apariciones conjuntas de los ítems. En la figura 7 se pueden apreciar las 20 reglas de asociación más fuertes según su confianza.

La aparición constante de productos de tipo fruta o verdura se debe a que la fracción del *dataset* utilizado corresponde a la temporada de verano, por lo que es comprensible este comportamiento. Esto indica que es importante realizar un análisis exploratorio previo a la *data* para entender, luego, los resultados obtenidos.

No se decidió por realizar más pruebas con diferentes parámetros porque estos dependen enteramente de las necesidades de negocio del *retailer*. Los resultados obtenidos de otra combinación de parámetros no son más ni menos correcta que la realizada en este trabajo. El punto clave para la obtención de los mejores resultados es dedicar el esfuerzo al entendimiento y tratamiento de la *data* de entrada y el estudio profundo de las necesidades de la organización.

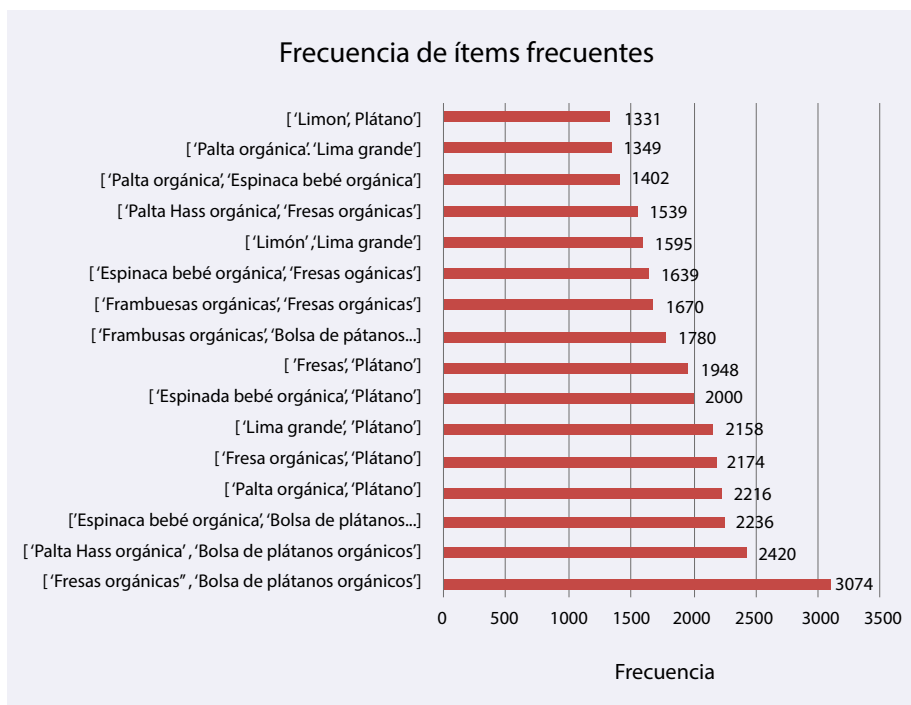


Figura 6. Frecuencia de ítems  
Elaboración propia

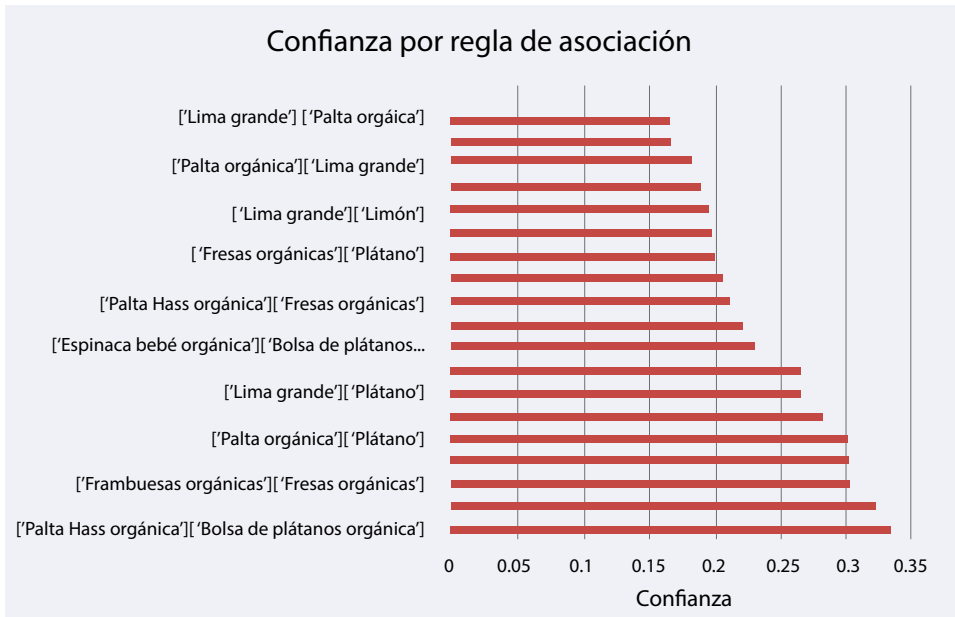


Figura 7. Confianza por regla de asociación  
Elaboración propia

Respecto al desempeño del modelo, se realizaron varias pruebas al tiempo de ejecución del modelo y sus componentes. Para esto se identificaron las partes que componen el modelo: el ajuste de la *data* al modelo (*model fit*), la generación de ítems frecuentes, la generación de reglas de asociación, la lectura o carga de la *data*, la preparación de la *data* para alimentar al modelo. En la tabla 1 se pueden observar los tiempos de ejecución para cada uno de estos componentes. El número de evaluaciones hace referencia a cuántas veces se ejecutará el componente correspondiente, mientras que las repeticiones significan cuántas veces esta evaluación será realizada. La salida de este ejercicio es el tiempo más bajo del total de repeticiones, mientras que el tiempo de ejecución es la división del tiempo total de repetición entre el número de evaluaciones. Es decir, para la primera fila, se ejecutó 5 veces el componente *model fit* con 10 evaluaciones cada una, el mínimo tiempo entre estas 5 repeticiones es el tiempo total de repetición (44,21 s), y este tiempo dividido entre el número de evaluaciones da el tiempo promedio (para uso referencial) de cada evaluación (4,42 s).

Es tentador calcular la media y la desviación estándar del vector resultado de las repeticiones. Sin embargo, esto no es muy útil. En un caso típico, el valor más bajo proporciona un límite inferior para la rapidez con la que la máquina pueda ejecutar el fragmento de código dado; los valores más altos en el vector de resultados generalmente no son causados por la variabilidad en la velocidad de Python, sino por otros procesos que interfieren con su precisión de tiempo. Por lo tanto, el mínimo del resultado es probablemente el único número que debería utilizarse como referencia al tiempo de ejecución del componente.

Tabla 1  
*Tiempo de ejecución de los componentes del modelo en segundos*

Tarea	Repeticiones	Número de evaluaciones	Tiempo de ejecución (s)	Tiempo total repetición (s)
<i>model fit</i>	5	10	4,421917761	44,21917761
<i>model fit</i>	10	10	4,375176091	43,75176091
<i>freq items</i>	5	10	0,000235505	0,002355054
<i>freq items</i>	10	10	0,000227227	0,002272273
<i>ass rules</i>	5	10	0,001781023	0,01781023
<i>ass rules</i>	10	10	0,001778434	0,017784339
<i>load data</i>	5	5	11,37215843	56,86079213
<i>load data</i>	5	10	11,8818465	118,818465
<i>prepare data</i>	5	3	8,014044508	24,04213352
<i>prepare data</i>	5	5	7,0520125	35,2600625

Elaboración propia

Si se separan los componentes por metodología, presentados en la tabla 1, en dos grupos, se tendría a *model fit*, *freq items* (ítems frecuentes) y *ass rules* (reglas de asociación) como un grupo y a *load data* y *prepare data* como otro. Este último es muy costoso en tiempo porque realiza la carga de la *data* (3 millones de registros para este trabajo) y las asociaciones necesarias. Para el primer grupo, es *model fit* el que toma mayor tiempo pues ajusta la *data* al modelo, el cual crea el árbol FP a partir de los parámetros establecidos, para posteriormente tanto *freq items* como *ass rules* realicen la minería de ese árbol. En este punto es donde se aprecian las ventajas de FP-Growth, ya que, una vez realizada la construcción del árbol, la minería de esta y la obtención de información como resultado toma milésimas de segundo para un conjunto de *data* real.

En cuanto al tiempo de ejecución variando la cantidad de registros a procesar, en las tablas 2 y 3 se pueden observar los resultados obtenidos. En la tabla 2 se realiza la misma dinámica que en la tabla 1. Se utilizó el módulo *Timeit* de Python para realizar la medición del tiempo. Como puede observarse, la diferencia de tiempo no es drástica entre las pruebas. Aunque el impacto de subir de 100 000 a 1 000 000 registros fue fuerte y el más notorio entre todos, el siguiente paso arriba (10 000 000) no tiene gran diferencia. La tabla 3 mide lo mismo, la diferencia radica que esta utiliza el módulo *Time* para medir el tiempo. Esa medición, a diferencia de la anterior, es afectada por los procesos que se ejecutan en segundo plano u otras tareas que el ordenador está realizando.

Tabla 2

*Tiempo de ejecución del ajuste del modelo en base a la cantidad de registros a procesar en segundos, calculado con el módulo Timeit*

TIMEIT.TIMEIT				
Cantidad de registros	Repeticiones	Número de evaluaciones	Tiempo de ejecución (s)	Tiempo total de repetición (s)
1000	5	10	3,535814589	35,35814589
10 000	5	10	3,461978596	34,61978596
100 000	5	10	3,606037736	36,06037736
1 000 000	5	10	4,492759976	44,92759976
10 000 000	5	10	4,421917761	44,21917761

Elaboración propia

Tabla 3

*Tiempo de ejecución del ajuste del modelo en base a la cantidad de registros a procesar en segundos, calculado con el módulo Time*

TIME.TIME	
Cantidad de registros	Tiempo de ejecución (s)
1000	3,628935814
10 000	3,599936247
100 000	3,862709284
1 000 000	4,239717007
10 000 000	5,348645687

Elaboración propia

## 6. CONCLUSIONES

Al ser una de las tendencias la generación de *data* y la necesidad de las organizaciones a obtener información y generar conocimientos de esta *data*, unos de los sectores con más oportunidades de minería de *data* es el sector *retail*. Cuando se habla de la minería de reglas de asociaciones siempre se ha tenido como referente al algoritmo Apriori, sin embargo, el algoritmo utilizado en este trabajo se presenta como una alternativa a las deficiencias del algoritmo clásico. FP-Growth propone la abstracción de la *data* en una estructura árbol y luego la minería de este, reduciendo así los costos en tiempo y recursos computacionales.

En este trabajo se realiza la implementación de FP-Growth para un conjunto de datos de una empresa de *retail* peruana perteneciente a los autoservicios. Se realiza una implementación



en paralelo haciendo uso del *framework* Apache Spark, el cual tiene como ventaja hacer procesamiento paralelo, con una limitante, todo fue construido sobre un solo ordenador, lo que no permite explotar las capacidades del procesamiento en clúster. Existen otras soluciones como la implementación en clústeres o en máquinas virtuales, asimismo existen soluciones novedosas implementadas en la nube como es el servicio de Amazon. Ellos brindan un ambiente de clústeres preparados para funcionar con el *framework* Spark.

Finalmente, hay que considerar que las necesidades del negocio y sus objetivos estratégicos juegan un papel dominante y crítico en la implementación de un modelo como el propuesto en este trabajo. Este análisis define los parámetros de funcionamiento del modelo y los resultados obtenidos tienen que ser relevantes y consistentes con la información del negocio.

## REFERENCIAS

- Agrawal, R., Imienlinski, T., y Swami, A. (1993). Mining association rules between sets of items in large databases. *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data*, 22(2), 207-216. doi:10.1145/170035.170072
- Bhandari, A., Gupta, A., y Das, D. (2015). Improvised Apriori algorithm using frequent pattern tree for real time applications in data mining. *Procedia Computer Science*, 46, 644-651. doi:10.1016/j.procs.2015.02.115
- Chunhua, J., y Dongjun, N. (2008). Distributed mining model and algorithm of association rules for Chain retail enterprise. *Proceedings - ISECS International Colloquium on Computing, Communication, Control, and Management, CCCM 2008*, 3, 235-239. doi:10.1109/CCCM.2008.129
- Di Fatta, G. (2019). Association rules and frequent patterns. En S. Ranganathan, K. Nakai, y C. Schonbach, (Eds.), *Encyclopedia of Bioinformatics and Computational Biology* (pp. 367-373). doi:10.1016/B978-0-12-809633-8.20333-6
- Fournier-Viger, P., Lin, J. C. W., Vo, B., Chi, T. T., Zhang, J., y Le, H. B. (2017). A survey of itemset mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 7(4), 1-41. doi:10.1002/widm.1207
- Galarreta Vásquez, J. (2016). Inducción de reglas de asociación de minería de datos en base de datos de entidad retail. *Rev. Ingeniería: Ciencia, Tecnología e Innovación*, 3(2).
- Girotra, M., Kanika, N., Saloni, M., y Neha, S. (2013). Comparative survey on association rule mining algorithms. *International Journal of Computer Applications*, 84(10), 975-8887. Recuperado de <https://pdfs.semanticscholar.org/08a7/a7c571159d14660a402a4460ee1f828c5fed.pdf>

- Griva, A., Bardaki, C., Pramatari, K., y Papakiriakopoulos, D. (2018). Retail business analytics: Customer visit segmentation using market basket data. *Expert Systems with Applications*, 100, 1-16. doi:10.1016/j.eswa.2018.01.029
- Han, J., Kamber, M., y Pei, J. (2011). *Data mining: Concepts and techniques*. doi:10.1016/C2009-0-61819-5
- Han, J., Pei, J., y Yin, Y. (2000). Mining frequent patterns without candidate generation. *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*, 1-12.
- Hussain, R. Z., y Srivatsa, S. K. (2014). A study of different association rule mining techniques. *International Journal of Computer Applications*, 108, 10-15.
- Kaur, M., y Kang, S. (2016). Market Basket Analysis: Identify the changing trends of market data using association rule mining. *Procedia. Procedia Computer Science*, 85, 78-85. doi:10.1016/j.procs.2016.05.180
- Tan, P., Steinbach, M., y Kumar, V. (2005). *Introduction to Data Mining. Discovering Knowledge in Data*. doi:10.1002/0471687545.ch1
- Xue, L., Wang, H., Liu, S., y Li, C. (2010). The application of data mining in online bookstore. *2010 International Conference on Machine Learning and Cybernetics*, 1294-1298. doi:10.1109/ICMLC.2010.5580892

# PÓSTERES

# COMPARACIÓN ENTRE REGRESIÓN LOGÍSTICA Y *RANDOM FOREST* PARA DETERMINACIÓN DE FACTORES DE VIOLENCIA DE PAREJA EN EL PERÚ

Ashley Mercedes Guerrero-Muguerza

La violencia de pareja es una problemática social que ha sido estudiada por diferentes investigadores para los factores que influyen en la ocurrencia de la misma, considerando diferentes entornos, tiempos y locaciones. El 68,2 % de mujeres han sido víctimas de violencia, y el 31,7 % fueron víctimas de agresión física en el Perú. La presente investigación propone nueve modelos basados en logística y *random forest* con las de *chi-square*, entropía y Gini, y tres sub escenarios de cinco, diez y veinte variables que utilizaron el *dataset* de denuncias registradas en el año 2016 del Ministerio de la Mujer. Se obtuvo el mejor resultado de cada subescenario, pero finalmente el mejor modelo fue el de veinte variables utilizando el *feature selection random forest (entropy)* y el modelo *random forest (Gini)*.

## Comparison Between Logistic Regression and Random Forest for Determining Factors of Intimate Partner Violence in Peru

Intimate partner violence is a social problem that has been studied by different researchers to determine the factors that influence its occurrence, considering different environments, moments and locations. Sixty-eight point two percent (68.2%) of women have been victims of violence and 31.7% have been victims of physical aggression in Peru. The present research proposes nine models based on logistic regression and random forest with variants such as chi-square, Entropy and Gini, and three subscenarios out of five, ten and twenty variables that used the dataset of complaints registered in 2016 at the Ministry of Women. The best result of each subscenario was obtained, but finally the best model was that of twenty variables which used the random forest "feature selection" (Entropy) and the random forest model (Gini).

# Comparación entre regresión logística y *random forest* para determinación de factores de violencia de pareja en el Perú

Ashly Mercedes Guerrero -Muguerza  
20150626@aloe.ulima.edu.pe

**RESUMEN:** La violencia de pareja es una problemática social que ha sido estudiada por diferentes investigadores para determinar los factores que influyen en la ocurrencia de la misma, considerando diferentes entornos, tiempos y locaciones. El 68,2 % de mujeres han sido víctimas de violencia, así como el 31,7 % lo fueron de agresión física en el Perú. La presente investigación propone nueve modelos basados en regresión logística y *random forest* con las variantes de *chi-square*, entropía y Gini, y tres subescenarios de cinco, diez y veinte variables que utilizaron el *dataset* de denuncias registradas en el año 2016 del Ministerio de la Mujer. Se obtuvo el mejor resultado de cada subescenario, pero finalmente el mejor modelo fue el de veinte variables utilizando el *feature selection random forest (entropy)* y el modelo *random forest (Gini)*.

## Introducción

Actualmente entre los diversos problemas sociales sobresale la violencia de pareja, el cual es un tipo de violencia de género, como el más grave. Este crimen es categorizado según la forma de violencia o acción cometida.

- Violencia física es la realización de lesiones, fisuras, contusiones, entre otros daños, a la pareja. El 31,7 % de mujeres han sufrido este tipo de agresión en el Perú.
- Violencia psicológica consiste en dominar y confinar a la víctima, en contra de su voluntad, y causarle daños psíquicos. El 64,2 % de mujeres peruanas han vivido situaciones que involucran este tipo de agresión.
- Violencia sexual ocurre cuando se realiza cualquier tipo de coacción con fines sexuales en contra de una persona sin su consentimiento. El 6,6 % de las mujeres en el Perú han reportado ser víctimas de esta agresión. (INEI, 2017)
- Violencia económica o patrimonial consiste en realizar un deterioro económico o financiero a la pareja (economía y patrimonio). El 68,2 % de mujeres peruanas han sufrido este tipo de agresión en el Perú. (INEI 2017)

## Materiales y métodos



### Preprocesamiento de datos

La limpieza de datos consiste en utilizar diversas técnicas para estandarizar la data, en otras palabras, eliminar variables identificadas con un determinado porcentaje de nulos o variables que poseen datos nulos, pero son fáciles para el modelo por lo que se deben utilizar técnicas para poder completarlos. En este caso se utilizó la moda para variables como la edad, los años, y la moda para variables categóricas.

### Identificación de factores influyentes

Se utilizó *feature selection* para obtener los factores que están más relacionados a la variable predictora la cual es la violencia física. Se realizaron tres técnicas, la primera fue *chi-square* la cual pertenece al grupo de los métodos de filtro (*filter method*), la segunda y tercera fueron el *random forest* con el criterio de división de nodos de Gini y entropía, respectivamente. Por cada técnica se realizaron diferentes subescenarios de cinco, diez y veinte variables; de los que se obtuvieron nueve escenarios diferentes.

### Datos

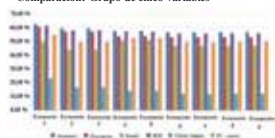
Los *datasets* utilizados se obtuvieron del Ministerio de la Mujer y Poblaciones Vulnerables. Este registro pertenece a las denuncias realizadas a nivel nacional, las cuales presentan información de las víctimas o personas que conocían a víctimas de violencia durante el año 2016. De las denuncias presentadas se seleccionaron, únicamente, las relacionadas con la violencia física de pareja.

## Resultados

En la tabla se pueden observar los nueve escenarios realizados:

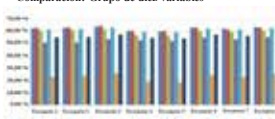
Escenario	Modelo	Subescenario	Resultado
1	Logistic Regression	Chi-square	Escenario 1
2	Logistic Regression	Entropy	Escenario 2
3	Logistic Regression	Gini	Escenario 3
4	Random Forest	Chi-square	Escenario 4
5	Random Forest	Entropy	Escenario 5
6	Random Forest	Gini	Escenario 6
7	Random Forest	Entropy	Escenario 7
8	Random Forest	Gini	Escenario 8
9	Random Forest	Entropy	Escenario 9

### Comparación: Grupo de cinco variables



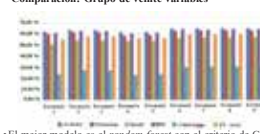
- El mejor modelo es la regresión logística con *chi-square* (escenario 1).
- ([2647, 974], [1528, 1499]).
- Cohen's kappa: 0,2297.
- Tomando como referencia la matriz de confusión, el valor de Cohen's kappa es el adecuado, debido al valor de los clasificados como violencia física.
- La regresión logística tiene una *performance* buena al utilizar *random forest*.
- Los escenarios 6 y 7 son los que poseen la más baja *performance*.

### Comparación: Grupo de diez variables



- El mejor modelo es el de la regresión logística con el *feature importance* utilizando el criterio Gini.
- ([2624, 997], [1418, 1609]).
- Cohen's Kappa: 0,2597.
- Tomando como referencia la matriz de confusión, el valor de Cohen's kappa es el adecuado, como resultado del valor de los clasificados como violencia física.
- El segundo mejor es el escenario 8, seguido de cerca por el escenario 6.
- Utilizar *chi-square* con *random forest* no ayuda en la *performance* del modelo.

### Comparación: Grupo de veinte variables



- El mejor modelo es el *random forest* con el criterio de Gini para la división de nodos y la selección de variables importantes con el criterio de *entropy* (escenario 9).
- ([2627, 994], [1264, 1763]).
- Cohen's kappa: 0,3047
- Tomando como referencia la matriz de confusión, el valor de Cohen's kappa es el adecuado, debido al valor de los clasificados como violencia física.
- El segundo mejor es el escenario 7, en el cual sus valores son parecidos al escenario seleccionado.
- El uso de *random forest* y la regresión logística con *chi-square* se encuentran entre los peores resultados.

- El modelo seleccionado fue el de las veinte variables, debido a que la mayoría de los factores obtenidos para dichos subescenarios se encontraron en la literatura revisada.
- Las variables más importantes del modelo son la edad de la víctima y del agresor, el tiempo de duración de la agresión, el nivel educativo de ambas partes y el número total de hijos. En comparación a los últimos cuatro factores listados, los cuales no son factores que tienen una gran importancia en el modelo, según los resultados obtenidos por el *random forest*.

## Conclusiones

El uso de la regresión logística, ya sea con *chi-square* o *random forest*, para la selección de características, es recomendable dados sus buenos resultados. Sin embargo, se debe utilizar cuando los factores independientes no sean de gran cantidad, pues se observó que durante la investigación, cuando la cantidad de factores era mayor, el modelo que ofrecía mejores resultados era el *random forest*.

Cabe recalcar que, para seleccionar el mejor modelo, no se debe utilizar el *accuracy* como única métrica para ver la *performance* del modelo, debido a que a) cuando los datos están desbalanceados esta métrica no aporta mucho en el resultado y b) utilizar diversas métricas conduce a una mejor comprensión del modelo utilizado.

## Referencias

- Abramsky, T., Watts, C. H., García-Moreno, C., Devries, K., Kiss, L., Ellsberg, M., ... y Heise, L. (2011). What factors are associated with recent intimate partner violence? Findings from the WHO multi-country study on women's health and domestic violence. *BMC Public Health*, 11(1), 109.
- Alves, L. G., Ribeiro, H. V., y Rodrigues, F. A. (2018). Crime prediction through urban metrics and statistical learning. *Physica A: Statistical Mechanics and its Applications*, 505, 435-443.
- Belgin, M., y Düzgün, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24-31.
- Berk, R. A., Sorenson, S. B., y Barnes, G. (2016). Forecasting domestic violence: A machine learning approach to help inform arraignment decisions. *Journal of Empirical Legal Studies*, 1(1), 94-115.
- Hsieh, T. C., Wang, Y. H., Hsieh, Y. S., Ke, J. T., Liu, C. K., y Chen, S. C. (2018). Measuring the unmeasurable - a study of domestic violence risk prediction and management. *Journal of Technology in Human Services*, 36(1), 56-68. doi:10.1080/1522835.2017.1417953
- Instituto Nacional de Estadística e Informática (INEI) (2017). *Perú: Indicadores de violencia familiar y sexual, 2000-2017*.
- Izmitli, G., Sommez, Y., y Sezik, M. (2014). Prediction of domestic violence against married women in southwestern Turkey. *International Journal of Gynecology and Obstetrics*, 127(3), 288-292.
- Kraanen, F., Vold, E., Scholing, A., y Emmelkamp, P. (2014). Prediction of intimate partner violence by type of substance use disorder. *Journal of Substance Abuse Treatment*, 46(4), 532-539.
- Sale, R., Neuner, F., Irti, V., y Catani, C. (2013). Prevalence and predictors of partner violence against women in the aftermath of war: A survey among couples in Northern Uganda. *Social Science & Medicine*, 86, 17-25.

## Agradecimientos

Agradezco a mis profesores, quienes en todo momento me alentarón y apoyaron en el desarrollo del presente trabajo; especialmente a los ingenieros Juan Gutiérrez, Vilma Romero, Rosario Guzmán y Pablo Rojas, y a la magister Rosa Millones.

## BLOCKCHAIN Y *SMART CONTRACT* PARA LA TRAZABILIDAD DE LAS DONACIONES

Rossy Espinoza / Carlos Ugaz

En la actualidad, una organización que administra donaciones tiene dificultades para controlar el flujo de ingreso y salida de donativos, dificultando la trazabilidad de estos. En esta investigación se propone un sistema basado en *blockchain* que garantice la trazabilidad de los donativos desde su origen hasta su destino, utilizando un protocolo de consenso *proof of work* y un contrato inteligente (*smart contract*) con funciones específicas.

### Blockchain and Smart Contract for Traceability of Donations

At present, an organization that administers donations has difficulties in controlling the inflow and outflow of such donations, hindering their traceability. This research proposes a blockchain-based system that guarantees the traceability of donations from their origin to their destination, using a proof of work consensus protocol and an intelligent contract with specific functions.

# BLOCKCHAIN Y SMART CONTRACT PARA LA TRAZABILIDAD DE LAS DONACIONES

Rossy Espinoza  
20131797@aloe.ulima.edu.pe

Carlos Ugaz  
20131349@aloe.ulima.edu.pe

**RESUMEN:** En la actualidad, las organizaciones que administran donaciones tienen inconvenientes para controlar el flujo de ingreso y salida de donativos, dificultando la trazabilidad de estos. En esta investigación se propone un sistema basado en *blockchain* que garantice dicha trazabilidad, desde su origen hasta su destino, utilizando un protocolo de consenso *proof of work* y un contrato inteligente (*smart contract*) con funciones específicas.

## INTRODUCCIÓN

- Falta de control de donaciones, pues ninguna organización de gestión de donaciones cuenta con un sistema exclusivo para el donante que le permita realizar la trazabilidad a sus donativos.
- El grupo Servir, de la Universidad de Lima, carece de un sistema de control de flujos.
- Se plantea la creación de un sistema basado en *blockchain* con la integración de un contrato inteligente (*smart contract*) que permita la implementación de validaciones y restricciones en la gestión de los donativos, para permitir la trazabilidad de estos.

## DIAGRAMA DE LA SOLUCIÓN



## MÉTODOS

### Blockchain



### Protocolo de consenso

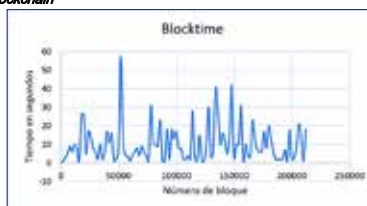


### Funciones del contrato Inteligente (*smart contract*)

- Crear usuario.
  - Obtener usuario.
- Crear donación.
  - Obtener donación.
- Enviar donación.

## RESULTADOS

### Blockchain



Variación del tiempo de generación de bloques

### Trazabilidad

Al momento de registrar una donación se crea una transacción en la que se almacena la clave pública del donante y la del grupo Servir, y al enviar la donación, la transacción que se efectúa contiene esa clave pública tanto de Servir como del beneficiario.

#	Transacción	Origen	Destino	Hash
1	Crear donación	0x2918867206881516028140779175d890086a77	0x2918867206881516028140779175d890086a77	0x543728d76a178e14884c290610089428a4d59842d95d198a799058ba
2	Enviar donación	0x57878725642c4c7077c74e48f482901418282447a	0x7994a605c11a668461304681c258826823a	0x57767676830c9878161c304681c258826823a300008826823a

Transacciones realizadas para verificar la trazabilidad

### Aplicativo

Responsable	Escenario	Resultado
Donante	Crear cuenta	✓
Donante	Iniciar sesión	✓
Grupo Servir	Crear donación	✓
Grupo Servir	Crear beneficiario	✓
Grupo Servir	Enviar donación	✓
Donante y Grupo Servir	Visualizar estado del donativo	✓

Funcionalidades del aplicativo

## CONCLUSIONES

- El sistema para la trazabilidad de donativos es aceptado por la encargada del grupo Servir.
- Se pudo demostrar que un sistema basado en *blockchain* con un contrato inteligente (*contract smart*) puede ser utilizado para la gestión de donativos, permitiendo validar la trazabilidad de estos.
- El contrato inteligente, según la evaluación por juicio de expertos, está a un nivel adecuado para la solución propuesta.
- Se demostró la compatibilidad del contrato inteligente (*contract smart*) con el sistema basado en otro *blockchain* existente.

## REFERENCIAS

Abdellatif, T., y Brousmiche, K.-L. (2018). Formal verification of smart contracts based on users and blockchain behaviors model. *2018 8th IFIP International Conference on New Technologies, Mobility and Security (NTMS)*. doi:10.1109/NTMS.2018.8328737

Gervais, A., Karamé, G. O., Wüst, K., Glykantzis, V., Ritzdorf, H., y Capkun, S. (2016). On the Security and Performance of Proof of Work Blockchains. *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (CCS16)*. Association for Computing Machinery, New York, 3-16. doi:10.1145/2976749.2978241

Nakamoto, S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System. Recuperado de <https://bitcoin.org/bitcoin.pdf>. doi:10.1.1.221.9986

## AGRADECIMIENTOS

Agradecemos a nuestras familias, al magister Daniel Cárdenas, al ingeniero Juan José Miranda, a la magister Natalie Gil y a Andrea Cueva, y a todas las personas que intervinieron y nos apoyaron en esta investigación.

## COMPARACIÓN DE MÉTODOS PARA CLASIFICAR COMENTARIOS DE LUGARES TURÍSTICOS POR MEDIO DE ANÁLISIS DE ASENTIMIENTO

Luis Guillermo Herrera-Sarmiento

Hoy en día los turistas luego de visitar algún destino, plasman sus experiencias como opiniones en diversas fuentes digitales, siendo información valiosa para empresas turísticas o relacionadas para identificar qué sitios son una oportunidad de mejora para los turistas durante la planificación de sus viajes. En esta investigación se propone la comparación de Support Vector Machine (SVM), Naïve Bayes (NB) y método propuesto basado en SVM y *chi square* como método de selección de características. La técnica híbrida propuesta obtuvo el mejor resultado, seguido de SVM y por último Naïve Bayes, cada una con 76,50 %, 67,53 % y 66,91 % de precisión, respectivamente.

### Comparison of Methods for Classifying Comments on Tourist Places by Sentiment Analysis

Nowadays, tourists express their experiences as opinions in various digital sources after visiting a destination, which is considered a valuable information for tourist companies or other related companies to identify which places are an opportunity for improvement, and for tourists when planning their trips. This research proposes the comparison of the support vector machine (SVM), naïve Bayes (NB), and a suggested method based on SVM and chi-square as a feature selection method. The proposed hybrid technique obtained the best result, followed by SVM and finally naïve Bayes, each with 76.50%, 67.53% and 66.91% accuracy, respectively.



# COMPARACIÓN DE MÉTODOS PARA CLASIFICAR COMENTARIOS DE LUGARES TURÍSTICOS POR MEDIO DE ANÁLISIS DE SENTIMIENTO

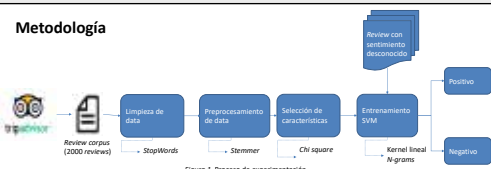
Luis Guillermo Herrera-Sarmiento  
20141885@aloe.ulima.edu.pe

**Resumen** Es común que los turistas, luego de visitar algún destino, plasmen sus experiencias y opiniones en diversas fuentes digitales, convirtiendo esa información en una herramienta valiosa para empresas turísticas o relacionadas porque así podrían identificar qué sitios son una oportunidad de mejora para los turistas durante la planificación de sus viajes. En esta investigación se propone la comparación de Support Vector Machine (SVM), Naive Bayes (NB) y el método propuesto basado en SVM y *chi square* como método de selección de características. La técnica híbrida propuesta obtuvo el mejor resultado, seguido de SVM y, por último, Naive Bayes, cada una con 76,50 %, 67,53 % y 66,91 % de precisión, respectivamente.

## Introducción

Se ha desarrollado una tendencia que va en aumento sobre el número de viajeros independientes o *backpackers* (Chi, Lo, Chu, y Lin, 2009). El Ministerio de Comercio Exterior y Turismo del Perú (MINCETUR) indica que el crecimiento del turismo en el país en el año 2018 fue de 9,6 % a comparación del año anterior y el Scotiabank proyecta un incremento del 10 % para el año 2019. En el Perú es una de las industrias más grandes, puesto que es una de las principales fuentes de la economía peruana. Por ello se debe tomar en consideración toda información buena o mala acerca de los puntos de interés para poder encontrar puntos de mejora al servicio ofrecido. Lin, Wu, Chen, Ku, y Chen (2014) mencionan que los viajeros que llegan a su destino no encuentran sitios turísticos agradables, lo que conlleva a que estos fomenten una mala imagen del sitio por medios digitales. El propósito de esta investigación es identificar el nivel de agrado de los turistas a partir de sus comentarios en un sitio web turístico y clasificarlo ya sea en negativo o en positivo, puesto que, como menciona Chi et al. (2009), un viaje para un turista es una actividad personal que busca objetivos.

## Metodología



**Recolección de datos**  
Se extrajeron *reviews* de TripAdvisor (Li y Yang, 2017), dado que estos poseen variables como fecha, título, *rating* y ciudad (Parikh, Kestur, Dharia, y Gotmare, 2018) mediante *web scraping* para desarrollar el conjunto de datos a trabajar. Solo se tomaron en consideración los cinco sitios turísticos de la ciudad de Lima, Perú, más comentados y con una antigüedad no mayor a cinco meses (2000 *reviews*).

**Limpieza de datos**  
Para la limpieza de la información se reconocieron los elementos que no pudieron ser interpretados dado que en el paso anterior algunos de los elementos extraídos los contenían como se aprecia en la figura 2.

*«Visitamos el lugar en compañía de mi familia en el mes de abril, el complejo arqueológico es hermoso e interesante, muestra la arquitectura preincaica. <br /><br /> El guía nos comentó acerca de la forma en que se construyó y cómo ha sobrevivido al paso del tiempo.<br /><br />»*

*«Visitamos el lugar en compañía de mi familia en el mes de abril, el complejo arqueológico es hermoso e interesante, muestra la arquitectura preincaica. El guía nos comentó acerca de la forma en que se construyó y cómo ha sobrevivido al paso del tiempo.»*

Figura 2. *Review* elementos sin *reg*

### Desarrollo del clasificador

Este se basa en el algoritmo propuesto que consiste en dos fases: categorización y clasificación, que a su vez está compuesto de una serie de pasos que usan la medición con TF-IDF (término frecuencia de documento de frecuencia inversa), el modelo de *n-grams* y el estadístico *chi square* como *features selection*:

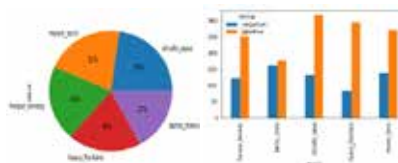
- Brindar pesos a las palabras mediante TF-IDF.
- Extracción de las características más relevantes con *chi square*.
- Clasificación del sentimiento del *review* con la técnica SVM.
- Comparación de técnicas Naive Bayes, SVM clásico y SVM propuesto.

## Resultados

Los resultados se plantean en dos instancias: 1) en el análisis de la información de TripAdvisor y 2) en el de la experimentación.

### Resultados de las vistas

En las siguientes imágenes se observa la proporción de *reviews* por lugar y la de sentimiento por lugar.



### Resultados de la experimentación

En la tabla se muestra la comparación de los niveles de *precision*, *recall*, *f1-score*, *accuracy*.

		<i>precision</i>	<i>recall</i>	<i>f1-score</i>	<i>support</i>
Prop	negative	0.62	0.46	0.53	114
	positive	0.8	0.89	0.84	283
	<b>avg/total</b>	<b>0.75</b>	<b>0.77</b>	<b>0.75</b>	<b>397</b>
SVM	negative	0.71	0.05	0.09	128
	positive	0.68	0.99	0.81	266
	<b>avg/total</b>	<b>0.69</b>	<b>0.68</b>	<b>0.58</b>	<b>394</b>
NB	negative	0.67	0.03	0.06	128
	positive	0.68	0.99	0.81	266
	<b>avg/total</b>	<b>0.68</b>	<b>0.68</b>	<b>0.56</b>	<b>394</b>

Técnica de clasificación de sentimiento	SVM	NB	Método propuesto
<i>Accuracy</i>	67,53 %	66,91 %	76,57 %

Como se observa en las tablas, la técnica híbrida propuesta obtuvo el mejor resultado: 76,50 % de precisión. Además, al momento de la predicción del sentimiento de un comentario fue correcta en un 75 % de las veces, identificó correctamente el 77 % el sentimiento y la exactitud de la clasificación fue del 75 %.

## Conclusiones

En esta investigación se aplicaron tres técnicas para la clasificación del texto: SVM propuesto, SVM clásico y Naive Bayes para la clasificación de comentarios turísticos. Detectando que las técnicas aplicadas pueden detectar y clasificar el sentimiento de los comentarios. En términos de exactitud, el modelo propuesto mostró la mejor *performance* alcanzando el 76,57 % en clasificación seguido del SVM clásico con 69,90 % y, por último, Naive Bayes con 66,91 % concordando con Thabtah, Eljinnai, Zamzeer, y Hadi (2009) que *chi square* como *feature selection* aumenta la exactitud del modelo. Este trabajo de investigación ha demostrado que es factible realizar la clasificación de manera automática e identificar qué lugares poseen el mayor ratio de reseñas.

## Referencias

Chi, T. H., Lo, H. H., Chu, Y. H., y Lin, W. C. (2009). A mobile tourism application model based on collective interactive genetic algorithms. *Proceedings of the 2009 Fourth International Conference on Computer Sciences and Convergence Information Technology (ICCCIT '09)*. IEEE Computer Society, 244-249. doi:10.1109/ICCCIT.2009.280

Li, J. B., y Yang, L. B. (2017). A Rule-Based Chinese Sentiment Mining System with Self-Expanding Dictionary - Taking TripAdvisor as an Example. *2017 IEEE 14th International Conference on e-Business Engineering (ICEBE)*, Shanghai, 238-242. doi:10.1109/ICEBE.2017.45

Lin, K. C., Wu, S. H., Chen, L. P., Ku, T., y Chen, G. D. (2014). Mining the user clusters on Facebook fan pages based on topic and sentiment analysis. *Proceedings of the 2014 IEEE 15th International Conference*

*on Information Reuse and Integration (IEEE IRI 2014)*, 627-632. doi:10.1109/IRI.2014.7051948

Parikh, V., Kestur, M., Dharia, D., y Gotmare, P. (2018). A tourist place recommendation and recognition system. *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, 218-222.

Thabtah, F., Eljinnai, M. A. H., Zamzeer, M., y Hadi, M. (2009). Naive Bayesian based on chi square to categorize arabic data. *Communications of the IBIMA*, 10, 158-163.

Ye, Q., Zhang, Z., y Law, R. (2009). Sentiment classification of online reviews to travel destinations by supervised machine learning approaches. *Expert Systems with Applications*, 36(5), 6527-6535. doi:10.1016/j.eswa.2008.07.035

## SISTEMA DE MONITOREO DE CALIDAD DE AGUA EN TANQUES ELEVADOS USANDO SENSORES IOT

Alejandro Javier Ortiz-Bazalar

A diferencia de otros recursos, el agua no puede ser substituida. Todos los seres vivos del planeta Tierra sobreviven gracias al agua, es así que asegurar la calidad del agua usando controles proactivos es urgente y necesario. Este trabajo muestra una nueva alternativa para el monitoreo de agua potable, aprovechando nuevas tecnologías como: el internet de las cosas y la computación en la nube.

### Water Quality Monitoring System in Elevated Tanks Using IoT Sensors

Unlike other resources, water cannot be substituted. Every life on earth survives thanks to water, so ensuring water quality using proactive controls is urgent and necessary. This work shows a new alternative for monitoring drinking water, taking advantage of new technologies such as the Internet of Things and cloud computing.

# Sistema de monitoreo de calidad de agua en tanques elevados usando sensores IoT

Alejandro Javier Ortiz-Bazalar  
20140948@aloe.ulima.edu.pe

**Resumen** A diferencia de otros recursos, el agua no puede ser substituida. Todos los seres vivos del planeta Tierra sobreviven gracias al agua, es así que asegurar su calidad usando controles proactivos es urgente y necesario. Este trabajo muestra una nueva alternativa para el monitoreo de agua potable, aprovechando las nuevas tecnologías como lo son: el internet de las cosas y la computación en la nube.

## Introducción

La presente investigación tiene como objetivo principal construir un sistema de medición que realice una evaluación cuantitativa mediante el monitoreo seleccionado de determinados parámetros fisicoquímicos del agua, que permita monitorear la calidad de agua en tanques elevados en hogares de la ciudad de Lima, Perú.

De esta manera, la finalidad es generar un diagnóstico acerca de la calidad del agua que se almacena diariamente en los tanques elevados y que las personas consumen en su día a día, asimismo se busca promover la implementación de medidas preventivas y correctivas que reduzcan el riesgo de transmisión de enfermedades hídricas que afecten la salud de las personas que consumen este elemento líquido.



## Metodología

### 1. Sistema de medición

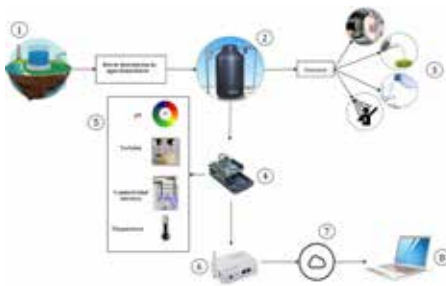
El sistema de medida consta de tres sensores: sensor de pH, CE y turbidez de agua. El microprocesador es el responsable de procesar los datos del sensor y la transmisión al concentrador de red (Alshattawi, 2017). Estos sensores van unidos al microcontrolador que lee los valores capturados por el sensor y estos se transmiten al concentrador de red, es decir, al gateway, que a su vez transmite los datos a la nube mediante conectividad wifi.

### 2. Infraestructura de red

Para la infraestructura de red se utiliza la modulación LoRaWAN. Esta es una red de baja potencia y de gran alcance (km), que tiene como objetivo eliminar los repetidores, aumentar la vida útil de la batería en los dispositivos, mejorar la capacidad de la red y admitir una gran cantidad de dispositivos (Khutsaone, Isong, y Abu-Mahfouz, 2017), por lo que su aplicación ayuda al ahorro de energía que es tan importante para la autonomía en este tipo de sistemas.

### 3. Interfaz para visualización de la información

Con la red en funcionamiento y enviando data en tiempo real, se muestra la información en la nube pública ThingSpeak. Se opta por realizar la interfaz de visualización de manera pública en pro de contribuir con el conocimiento de los usuarios sobre la calidad del agua a la que tienen acceso.



## Referencias

Alshattawi, S. K. (2017). Smart water distribution management system architecture based on internet of things and cloud computing. *2017 International Conference on New Trends in Computing Sciences (ICTCS)*, 289-294. doi:10.1109/ICTCS.2017.31

Khutsaone, O., Isong, B., y Abu-Mahfouz, A. M. (2017). IoT devices and applications based on LoRa/LoRaWAN. *The 43rd IEEE Conference of Industrial Electronic Society*, 6107-6112. doi:10.1109/IECON.2017.8217061

## Resultados

El sistema propuesto se testeó en un ambiente preparado, permitiendo medir los parámetros de calidad del agua elegidos según la investigación, se midieron el pH, la turbidez y la conductividad eléctrica del agua potable del tanque elevado seleccionado, corroborando así el correcto funcionamiento de los sensores, así como su calibración y programación (figura 1)



Figura 1. Calibración de los sensores y pruebas del sistema propuesto

De igual manera, en la parte de conexiones, el sistema de medición (Arduino, shield LoRa y sensores) con el gateway fue exitoso, pues se logró realizar la conexión, lo que permitió que el gateway pudiera enviar la trama obtenida hacia internet. La visualización en internet se dio por medio de la nube pública ThingSpeak, en ella se observó la variación cada ciertos minutos de los parámetros que estaban siendo medidos en tiempo real (figura. 2).



Figura 2. Información reflejada en la nube ThingSpeak

## Conclusiones

- Es posible la construcción de un sistema *low cost* confiable para el monitoreo de agua potable, pues los sensores usados para la obtención de parámetros han sido probados mediante calibración y tienen un precio más bajo respecto a los sensores profesionales, además su comportamiento mantiene un intervalo de confianza aceptable con relación a los equipos profesionales. Esta característica lo hace escalable y de mejor alcance para las personas, instituciones, etc., que quieran utilizar este sistema para el monitoreo de calidad del agua.
- La tecnología que se utilizó para la conexión, la modulación LoRaWAN, funciona bien para este tipo de sistemas por las ventajas expuestas en este trabajo, sin esta conexión el área de monitoreo quedaría bastante limitada.
- Se logró una autonomía del sistema con una estructura simple y de fácil entendimiento, así como una interfaz sencilla y eficiente al momento de mostrar la información.

## Agradecimientos

Agradezco a mi asesor, el profesor Carlos Torres, por el apoyo brindado en todo momento para el desarrollo de esta investigación. Igualmente, al profesor Javier Quino, quien me ayudó en lo referente a la calidad del agua y calibración de los sensores, que incluyó, entre otros, la preparación de soluciones estándar y el uso del equipo profesional del Laboratorio de Docimasia de la Universidad de Lima, y a cada una de las personas que de una u otra manera aportaron para que este trabajo se viese realizado.

## ANÁLISIS COMPARATIVO DE TÉCNICAS DE ESTEGANÁLISIS EN IMÁGENES DIGITALES LSB

Luis Sifuentes-Villarroel

En la presente investigación se evaluaron dos métodos de esteganálisis frente a imágenes LSB, comparando su efectividad en la detección de imágenes esteganográficas y del porcentaje que abarca el mensaje oculto en la imagen portadora (porcentaje de embebido) respecto a la totalidad de la imagen.

### Comparative Analysis of Steganalysis Techniques in LSB Digital Images

In the present research, two methods of steganalysis were evaluated against LSB images, comparing their effectiveness in detecting steganographic images and the percentage covered by the hidden message in the carrier image (percentage of embedding) with respect to the entire image.

# Análisis comparativo de técnicas de esteganálisis en imágenes digitales LSB

Luis Sifuentes-Villaruel  
20101062@aloe.ulima.edu.pe

**Resumen** En la presente investigación se evaluaron dos métodos de esteganálisis frente a imágenes LSB comparando su efectividad en la detección de imágenes esteganográficas y de la proporción que abarca el mensaje oculto en la imagen portadora (porcentaje de embebido) respecto a la totalidad de la imagen.

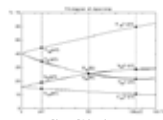
## Introducción

- La pornografía infantil está disponible en Internet en proporciones epidémicas, y los investigadores en línea están haciendo todo lo posible para indagar, detener y procesar a estos depredadores sexuales. Para lograr su cometido, dichos individuos utilizan una herramienta disponible en la web y de libre acceso llamada *esteganografía*, la cual usan para el ocultamiento y tráfico de este tipo de contenido ilícito.
- Hay más de 4000 sitios dedicados a la esteganografía, lo que les facilita ocultar una imagen de pornografía infantil dentro de otra aparentemente inocente.
- Con el pertinente conocimiento de esta herramienta (esteganografía), tanto los *software* comúnmente utilizados como las técnicas más frecuentes, un agente de la ley puede ser capaz de identificar de manera exitosa la posesión y transmisión de este tipo de contenido que normalmente pasaría desapercibido.



## Marco teórico

- La *esteganografía* consiste en técnicas que permiten ocultar mensajes u objetos, dentro de otros, llamados *portadores*, de modo que no se perciba su existencia.
- El esteganálisis es el arte de detectar los mensajes ocultos dentro de un estego-objeto (imagen oculta dentro de otra).
- El esteganálisis *RS (regular and singular)* consiste en estimar las cuatro curvas del diagrama RS y calcular su intersección mediante extrapolación. La forma general de las cuatro curvas en el diagrama varía con la imagen de portada, casi perfectamente lineal a la curva.



El eje *x* es el porcentaje de píxeles con LSB invertidos, el eje *y* es el número relativo de grupos regulares y singulares con máscaras *M* y  $-M = [0\ 1\ 1\ 0]$

- *Esteganálisis SPA (sample pair analysis)*. El método SPA rastrea los conjuntos múltiples de pares de muestras antes y después de la incrustación LSB, y utiliza las relaciones entre conjuntos múltiples para resolver la longitud de los mensajes incrustados.

## Resultados



Para el análisis de ambas técnicas se seleccionó la imagen Lenna en formato BMP. En primer lugar, se evaluó la técnica SPA con la imagen tanto a color como en escala de grises. Para el primer caso se mostraron los siguientes resultados:

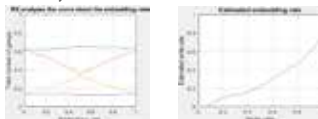
Señal	Tipo imagen	Ratio de embebido
R	Stego	0,095361
G	Stego	0,092576
B	Stego	0,125577

Los resultados mostraron el porcentaje de embebido estimado por cada frecuencia que compone el RGB siendo superior al 5 % de toda la imagen que es imagen esteganográfica. El mismo análisis se realizó para la misma imagen, pero en escala de grises.

Señal	Tipo imagen	Ratio de embebido
G	Stego	0,289529

Para el segundo caso, a diferencia de la imagen a color, para la imagen en escala de grises solo se mostró una frecuencia y el porcentaje de embebido incrementó al estar concentrado en un solo plano.

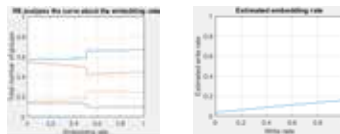
En segundo lugar, se tomó como primera muestra la imagen Lenna en escala de grises obteniéndose la gráfica RS que muestra la asociación de los conjuntos regulares y singulares generados en el proceso. Posteriormente, se realizaron diversos embebidos en la imagen que iban del 0 al 100 % y estos fueron contrastados con los embebidos estimados por la técnica.



Adicionalmente, se realizó el esteganálisis (detección de la existencia de una imagen esteganográfica) en una imagen con una mayor complejidad para identificar la efectividad de la detección de los datos embebidos en un plano más amplio.



Finalmente, los datos de este embebido mostraron que el porcentaje estimado era mínimo a causa de la alta dispersión permitida por esta imagen debido a su dimensión mayor y gama de colores.



## Conclusiones

- Si bien la técnica SPA puede detectar de manera confiable imágenes ocultas con tasas de bits de más de 0,05 y de igual forma funciona bastante bien con tasas de bits más bajas.
- El esteganálisis RS es notablemente eficaz; para evitar la detección, una incrustación de LSB tendría que ser rotada a menos de 0,05 bits por píxel en la imagen de portada, lo que significa que menos del 5 % de los píxeles han modificado su LSB.
- La robustez de la implementación de técnicas esteganográficas puede mejorar significativamente si las estego-imágenes (imagen que oculta otra dentro suya) se encuentran en escala de grises debido a que el ruido se vuelve más imperceptible por la baja variabilidad de tonalidades en este tipo de imágenes.

## Referencias

Fridrich, J., Goljan, M., y Du, R. (2001). Detecting LSB steganography in color, and gray-scale images. *IEEE Multimedia*, 8(4), 22-28. doi:10.1109/93.959097

Kessler, G. C. (2004). Steganography: implications for the prosecutor and computer forensics examiner. *American Prosecutors Research Institute Child Sexual Exploitation Program*.

Raju, K. B., Venugopal, K. R., y Patnaik, L. M. (2004). A Secure steganographic algorithm using LSB, DCT and image compression on raw images [technical report]. Bangalore: Department of Computer Science and Engineering, University Visvesvaraya College of Engineering, Bangalore University.

Xiangyang, L., Fenlin, L., y Peizhong, L. (2007). A LSB Steganography Approach against pixels sample pairs steganalysis. *International Journal of Innovative Computing, Information and Control*, 3(3), 575-588.

## Agradecimientos

Agradezco a mis padres y a mi hermana, por su constante apoyo y motivación durante mis estudios; al profesor del curso Seminario de Tesis 2, Daniel Cárdenas; y en especial a mi asesor, el profesor Carlos Torres, ya que sin su ayuda este trabajo no se hubiese podido realizar.

## SISTEMA DE INCREMENTO DE VOCABULARIO PARA LA MEJORA DE LA COMPRESIÓN LECTORA EN PRIMARIA CON AYUDA DE REALIDAD AUMENTADA

Guiabella Tania Urday-Ibarra

El Perú presenta resultados insatisfactorios en las evaluaciones internacionales y nacionales en el área de lectura. En los resultados del Programa para la Evaluación Internacional de Alumnos (PISA) del año 2015, el Perú obtuvo 398 puntos, casi 100 puntos por debajo de la media. El país se posicionó en el puesto 64 de los 72 participantes. En la evaluación nacional ECE, a medida que los grados avanzan el porcentaje de cumplimiento de objetivos académicos van decayendo. Pasa de 46,4 %, 31,4 % y 14,3 %, en el segundo y cuarto grado de primaria, y segundo de secundaria, respectivamente. El factor por acatar dentro del problema de la baja comprensión lectora es el vocabulario reducido.

### A System to Improve Vocabulary for Enhancing Reading Comprehension in Primary School with the Help of Augmented Reality

Peru has achieved unsatisfactory results in international and national reading assessments. In the 2015 Programme for International Student Assessment (PISA), Peru obtained 398 points, almost 100 points below the average. The country ranked 64 out of 72 participants. In the national ECE assessment, the higher the school grade, the less the achievement percentage of academic goals. Said percentage accounts for 46.4% and 31.4% in the 2nd and 4th grade of primary school, respectively, and 14.3% in the 2nd grade of secondary school. The factor that causes this problem is a poor vocabulary.

# Sistema de incremento de vocabulario para la mejora de la comprensión lectora en primaria con ayuda de realidad aumentada

Gianella Tania Urday-Ibarra

20141367@aloe.ulima.edu.pe

## Introducción

El Perú presenta resultados insatisfactorios en las evaluaciones internacionales y nacionales en el área de lectura. En los resultados del Programa para la Evaluación Internacional de Alumnos (PISA) del año 2015, el Perú obtuvo 398 puntos, casi 100 puntos por debajo de la media. El país se posicionó en el puesto 64 de los 72 participantes.

Así mismo, el Ministerio de Educación a nivel nacional también evalúa anualmente los logros de aprendizaje mediante la Evaluación Censal de Estudiantes (ECE), en sus resultados a medida que los grados avanzan el porcentaje de cumplimiento de objetivos académicos va decayendo. Pasa de un 46,4 % a 31,4 % a 14,3 %, en el segundo y cuarto grado de primaria y segundo de secundaria, respectivamente.

Uno de los factores destacados dentro del problema de la baja comprensión lectora es el vocabulario reducido. Según Defior (1996), el alumno tiene dificultad de entender lo que quiere transmitir el texto, si no comprende lo que significa.

## Objetivo

Desarrollar una aplicación basada en realidad aumentada para enriquecer el vocabulario que contribuya a la mejora de la comprensión lectora en niños de segundo grado de primaria.

## Resultados

Para la investigación participaron 22 alumnos de segundo grado de primaria. Los alumnos fueron divididos en dos grupos, de control y experimental. El grupo de control no tuvo acceso a la aplicación; mientras que el experimental, sí. Ambos grupos fueron evaluados con las mismas preguntas y sus resultados fueron comparados.

Promedio grupo de control	Promedio grupo experimental
5,5	5,8
Experimental sin app	Experimental con app
5	5,8

Referencias:  
 Azuma, R. T. (2017, June). Making augmented reality a reality: the Apple/iOS and Optical See-through AR. Inaugural Meeting (pp. 174-175). Optical Society of America.  
 MINEDU. (2016). Evaluación PISA 2015: Resumen ejecutivo. Recuperado del sitio del Ministerio de Educación. <http://www.minedu.gob.pe/wp-content/uploads/2016/12/presentacion-verb-PISA.pdf>  
 MINEDU. (2017). ¿Cuánto aprenden nuestros estudiantes? Resultados de la ECE 2014. Recuperado del sitio del Ministerio de Educación <http://www.minedu.gob.pe/wp-content/uploads/2017/04/Resultados-Evaluacion-2014.pdf>  
 Ceballos-Romero, J., Lora-Silvestre, J. J., Moreno-Martínez, X. M., Bascos-Ovaza, I., y López-Monera, E. (2010). Realidad aumentada y educación: Innovación en contextos formativos. Barcelona: Octaedro.  
 Defior, S. (1996). Una clasificación de los errores cometidos en la comprensión de las actividades de comprensión y algunas ideas para su mejor lectura y aprendizaje. (387-3), 49-61.  
 Martín-Alonso, R. M., Bascos-Soriano, G. (2017). La realidad aumentada aplicada al aprendizaje en procesos con sistemas de Ovario: un estudio exploratorio. Anales Latino de Comunicación Social, 72, 157-176.  
 Pineda-Torres, C. (2012). Realidad aumentada y educación: análisis de experiencias prácticas. Págs. 30. Revista de Pedagogía y Educación, 44, 35-50.  
 Tabón-Ardáiz, A. (2005). Comprensión lectora y procesos psicológicos. *Chilote*, 13(1), 49-61.

## Metodología

Realidad aumentada es una tecnología que permite que los objetos virtuales coexistan en la realidad física, al respecto, Prendes (2015) aclara que no tiene como propósito reemplazarla.

Azuma (2017) presenta tres características para que la realidad aumentada sea llamada como tal:

- La combinación de objetos reales y virtuales.
- Que sea en tiempo real.
- Que sea en 3D.



## Conclusiones

- La aplicación fue utilizada con éxito en un colegio privado con 22 alumnos de segundo grado de primaria.
- Se obtuvo una mejora en la comprensión lectora con una diferencia de 0,3 puntos entre los promedios del grupo de control y el grupo experimental.

## APLICACIÓN EN REALIDAD VIRTUAL PARA CONCIENTIZACIÓN DEL PELIGRO DE SOBREEXPOSICIÓN A LA RADIACIÓN UV

Edgard Javier Hernán Vargas-Solís

La presente investigación busca concientizar a la población acerca de la elevada radiación UV registrada en nuestro país, a partir del desarrollo de una aplicación en realidad virtual. Este estudio evaluará el desempeño de dicha aplicación a través de un cuestionario.

## A Virtual Reality Application to Raise Awareness on the Danger of Overexposure to UV Radiation

This research seeks to raise public awareness on high UV radiation levels recorded in our country. A virtual reality application was developed from a study and the performance of such application was evaluated through a questionnaire.



# Aplicación en realidad virtual para concientizar del peligro de sobreexposición a la radiación UV

Edgard Javier Hernán Vargas-Solís

20133252@aloe.ulima.edu.pe

## Resumen

La presente investigación busca concientizar a la población acerca de la elevada radiación UV registrada en nuestro país, a partir del desarrollo de una aplicación en realidad virtual. Este estudio evaluará el desempeño de dicha aplicación a través de un cuestionario.

## Introducción

- Según Lozano *et al.* (2016), los niveles de radiación UV en Lima, en casi todas las estaciones del año, son elevados.
- La población es indiferente al índice UV, al no percibirse del nivel de radiación UV al que se exponen en el día a día (Yaipen, Peralta, Rojas, Vivar, Herrera, y Diaz, 2019).

Tabla 1

Escala del índice UV utilizada por el SENAMHI

Valor del índice UV	Nivel de riesgo
1-2	Mínimo
3-5	Bajo
6-8	Moderado
9-11	Alto
12-14	Muy alto
14+	Extremo

Fuente: Lozano y Osón (2016)

## Metodología



Figura 1. Funcionamiento del sistema en realidad virtual

- Se desarrolló una aplicación de realidad virtual empleando Unity3D para Android que permite ver los efectos de la radiación en la piel, diagnosticar el tiempo mínimo antes de obtener una quemadura y recomendar el factor de protección (bloqueador) en base al índice UV del día para el tipo de piel según la escala de Fitzpatrick.
- Para el estudio, se reclutaron a 21 participantes (Schutte y Stilianović, 2017), quienes emplearon la aplicación desarrollada y respondieron a un cuestionario con las siguientes secciones: comportamiento dirigido a la salud, educación, usabilidad y atractivo.
- Antes de realizar la simulación, se entregaron a los participantes la sección de comportamiento dirigido a la salud del cuestionario, lo cual permitió evaluar sus conocimientos con respecto a la radiación UV.
- Finalmente, los participantes realizaron la simulación y se les dieron las secciones restantes del cuestionario para evaluar la aplicación desarrollada en los apartados de educación, usabilidad y atractivo.

## Resultados

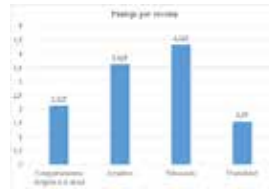


Figura 3. Resultados del cuestionario

- La sección de comportamiento dirigido a la salud obtuvo una media de 2,125 y un  $p$  valor de  $< 0,001$ .
- La sección de atractivo obtuvo una media de 3,635 y un  $p$  valor de 0,003.
- La sección de educación obtuvo una media de 4,325 y un  $p$  valor de  $< 0,001$ .
- La sección de atractivo obtuvo una media de 1,55 y un  $p$  valor de  $< 0,001$ .



Figura 2. Escenas de la aplicación en realidad virtual

## Conclusiones

- Los participantes tienen un conocimiento nulo o escaso acerca de la radiación UV, además de no poseer hábitos de cuidado de la piel.
- El entorno en realidad virtual logró satisfacer las exigencias de los usuarios, siendo lo suficientemente atractivo y absorbente, generándoles una experiencia agradable durante la simulación.
- La aplicación fue fácil de usar, lo cual permitió una simulación fluida y poco frustrante para los usuarios.
- La aplicación logró enseñar a los participantes sobre lo dañina que puede ser la radiación UV y cómo se podría evitar un daño significativo a la piel por medio del uso adecuado del bloqueador solar.

## Referencias

Yaipen-Salazar, C., Peralta-Palma, E., Rojas-Cubas, H., Vivar-Chao, A., Herrera-Toscani, A., y Diaz-Vélez, C. (2019). Conocimiento y actitudes de medidas de protección solar en trabajadores agrícolas. Tumbán mayo-junio 2014. *Revista del Cuerpo Médico del HNA44*, 8(4), 207-211.

Instituto Nacional de Enfermedades Neoplásicas (INEN) (2016). Manual de prevención del cáncer de piel inducido por la exposición prolongada a la radiación ultravioleta (RUV) [documento técnico]. Recuperado de [https://portal.inen.sld.pe/wp-content/uploads/2019/10/Cancer-de-piel-2018-op2\\_final.pdf](https://portal.inen.sld.pe/wp-content/uploads/2019/10/Cancer-de-piel-2018-op2_final.pdf)

Fitzpatrick, T. B. (1988). The validity and practicality of sun-reactive skin types I through VI. *Archives of Dermatology*, 124(6), 869-871. doi:10.1001/archderm.1988.0167006015008

Schutte, N. S., y Stilianović, E. J. (2017). Facilitating empathy through virtual reality. *Motivation and Emotion*, 41(6), 708-712.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 22(140), 5-55.

Lozano L., Rodríguez A., Sánchez O. (2016). Pronóstico con cobertura nacional del índice de radiación solar ultravioleta. SENAMHI. Recuperado de [https://www.senamhi.gob.pe/pdf/estudios/meteo\\_2016\\_pro\\_cob\\_nac\\_ind\\_rad\\_sol\\_ult.pdf](https://www.senamhi.gob.pe/pdf/estudios/meteo_2016_pro_cob_nac_ind_rad_sol_ult.pdf)

## IDENTIFICACIÓN DE MICROEXPRESIONES FACIALES DURANTE EL PROCESO DE SELECCIÓN DE PERSONAL

Luis Villanueva

¿Cómo saber si la información brindada por un postulante durante un proceso de selección de personal es verídica? Paul Ekman, psicólogo pionero en el estudio de las expresiones faciales, vincula las mentiras con las microexpresiones, un tipo de gesto facial sutil, corto y que representa emociones que una persona intenta reprimir. Esta investigación propone el uso de Local Binary Patterns y Support Vector Machines para detectar microexpresiones y clasificarlas a través de la implementación de un sistema que fue probado en la base de datos SAMM y en sujetos de prueba.

### Identifying Facial Microexpressions During a Personnel Selection Process

How to know if the information provided by an applicant during a personnel selection process is true? Paul Ekman, a pioneer psychologist in the study of facial expressions, links lies with microexpressions, a type of subtle, short facial gesture that represents emotions that a person tries to repress. This research proposes the use of “local binary patterns” and “support vector machines” to detect microexpressions and classify them through the implementation of a system evaluated in the SAMM database and test subjects.

# Identificación de microexpresiones faciales durante el proceso de selección de personal

Luis Villanueva  
20141468@aloe.ulima.edu.pe

**Resumen** ¿Cómo saber si la información brindada por un postulante durante el proceso de selección de personal es verídica? Paul Ekman, psicólogo pionero en el estudio de las expresiones faciales, vincula las mentiras con las microexpresiones, un tipo de gesto facial sutil, corto y que representa emociones que una persona intenta reprimir. Esta investigación propone el uso de Local Binary Patterns y Support Vector Machines para detectar microexpresiones y clasificarlas a través de la implementación de un sistema que fue probado en la base de datos SAMM y en sujetos de prueba.

## Introducción

- Manpower en 2018: de 40 000 empresas en 43 países, el 67 % admite que es difícil contratar porque no se encuentra a la persona ideal.
- RR.HH. identifica a potenciales candidatos para vacantes necesarias, buscando idoneidad.
- Una mala contratación generará gastos futuros como la repubicación del aviso, reclutamiento, evaluación y selección.
- Un error costoso es la mentira, en postulantes y empleados.
- Consecuencias del fraude: ética y efectos adversos a los objetivos del empleador.

## Conceptos y métodos

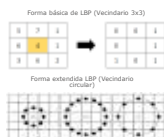
### Microexpresiones

- Sútiles
- Duran  $\frac{1}{4}$  s
- Emociones inconscientes
- Músculos fidedignos



### Local binary patterns

- Comparación de intensidades de píxeles en un vecindario.
- Se generan números binarios que se convierten en decimales para generar histogramas.

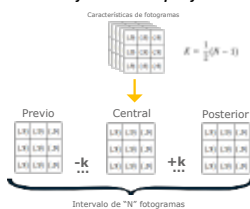


## Experimentación

### Extracción de características



### Identificación del Apex frame



$$Promedia = \frac{Previo + Posterior}{2} \leftrightarrow Promedia_{(N)} \\ Distancia = X^T(Promedia, Central) \leftrightarrow Distancia_{(N)} \\ D = Med_{(Distancias)} \leftrightarrow D_{(N)} \\ Y = \frac{\sum_{i=1}^N D_i}{3} \quad X = V - \frac{1}{2}(V_{+1} + V_{-1})$$

### Clasificación de microexpresiones



## Resultados

### En la BD SAMM

	SAMM
Detección	64 %
Clasificación	55 %

### En sujetos de prueba



	Sujetos de prueba
Clasificación	35 %

## Conclusiones

- La etapa de detección tuvo resultados aceptables, mientras que la de clasificación aún debe ser mejorada.
- Es necesario aumentar la cantidad de muestras en las bases de datos disponibles o tomarse como referencia bases de datos de macroexpresiones faciales.
- Investigaciones multidisciplinares futuras podrían incluir perfiles de comportamiento de personas para hacer un análisis más preciso.

## Referencias

Davison, A. K., Lansley, C., Costen, N., Tan, K., y Yap, M. H. (2018). SAMM: a spontaneous micro-facial movement dataset. *IEEE Transactions on Affective Computing*, 9(1), 116-129. doi:10.1109/TAFFC.2016.2573832

Davison, A., Merghani, W., Lansley, C., Ng, C.-C., y Yap, M. H. (2018). Objective micro-facial movement detection using FACS-Based Regions and Baseline Evaluation. *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, 642-649. doi:10.1109/FG.2018.00101

Moolanen, A., Zhao, G., y Pietikäinen, M. (2014). Spotting rapid facial movements from videos using appearance-based feature difference analysis. *2014 22nd International Conference on Pattern Recognition, IEEE*, 1722-1727. doi:10.1109/ICPR.2014.303

## Agradecimiento

Agradezco el apoyo de todos los profesores y alumnos que colaboraron en el desarrollo de esta investigación, especialmente a mi asesor, el profesor José Taquia Gutiérrez. Asimismo, extendiendo mi gratitud hacia los creadores de las bases de datos SAMM y CASME, quienes me permitieron utilizar sus imágenes y videos en la experimentación llevada a cabo.

## HONEYCLOUD - COMBINING RESEARCH AND TEACHING IN A PROJECT FOR THE DIGITALIZATION OF BEEKEEPING

Alexander Hilgarth / Michael Dorin / Sergio Montenegro

The HONEYCLOUD project is a research project of the University of Würzburg in the field of bioeconomics. Within the scope of technology transfer, tools and methods that were originally developed for aerospace applications are now to be made available for precision agriculture. In particular, the field of precision beekeeping was selected for this purpose. An IT infrastructure to support the honeybees and the work of beekeeping is to be developed. The aim is to ensure that the electronics remain unobtrusive and that the beehive does not become a switch cabinet. This results in an experimental test setup that can be used in the field of teaching Internet-of-Things (IoT) systems. In this way, hardware and software components are created in the project and used directly in teaching.

### HONEYCLOUD - Combinando investigación y docencia en un proyecto para la digitalización de la apicultura

El proyecto HONEYCLOUD es un proyecto de investigación de la Universidad de Würzburg en el campo de la bioeconomía. Dentro del alcance de una transferencia de tecnología, las herramientas y métodos que se desarrollaron originalmente para aplicaciones aeroespaciales ahora estarán disponibles para la agricultura de precisión. En particular, se seleccionó el campo de la apicultura de precisión para este propósito. Se debe desarrollar una infraestructura de TI para apoyar a las abejas y el trabajo de la apicultura. El objetivo es asegurar que la electrónica permanezca discreta y que la colmena no se convierta en un armario de distribución. Esto da como resultado una configuración de prueba experimental que se puede utilizar en el campo de la enseñanza de sistemas de Internet de las cosas (IoT). De esta forma, se crean componentes de *hardware* y *software* en el proyecto que se utilizan directamente en la enseñanza.

# HONEYCLOUD - Combining research and teaching in a project for the digitalization of beekeeping

Alexander Hilgarth, Michael Dorin, Sergio Montenegro

alexander.hilgarth@uni-wuerzburg.de, mike.dorin@stthomas.edu, sergio.montenegro@uni-wuerzburg.de

The HONEYCLOUD project is a research project of the University of Würzburg in the field of bioeconomics. Within the scope of technology transfer, tools and methods that were originally developed for aerospace applications are now to be made available for precision agriculture. In particular, the field of precision beekeeping was selected for this purpose. An IT infrastructure to support the honeybees and the work the beekeepers is to be developed. The aim is to ensure that the electronics remain unobtrusive and that the beehive does not become a switch cabinet. This results in an experimental test setup that can be used in the field of teaching Internet-of-Things (IoT) systems. In this way, hardware and software components are created in the project which are used directly in teaching.

## Introduction

Since the second half of the 19th century, the combination of research and teaching has been a practiced concept at German universities. In this tradition, students of the Chair of Aerospace Information Technology at the Julius Maximilian University of Würzburg are involved in real research projects. They are given the opportunity to work on the projects as student assistants or as part of their final thesis to acquire knowledge and to make concrete contributions. The same applies to the current HONEYCLOUD project, which will research whether a system for beehive monitoring could be implemented in practice.

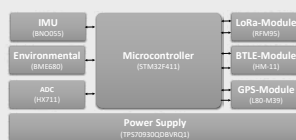
In terms of subject matter, the project is concerned with digitalization—i.e., the transformation of processes previously carried out analogously in the real world into a digital and networked mode of operation. This will be done by applying methods and tools from the aerospace domain to the field of bioeconomics. At first glance, the two domains appear to be far from each other. If compared, many differences can be formulated as opposites. For a manufacturer, space technology is a "market of one"—i.e., a market with a single customer and a single product specifically adapted for him/her. However, bioeconomy obviously affects all people and refers to a mass market. On closer examination, there are concrete similarities. In both areas, sensor data must be reliably recorded, forwarded, stored and evaluated in order to be able to regulate a system if necessary. The systems used should be small, light and energy-efficient.

## Materials and methods

The main tool developed at the department for the purpose of satellite programming is RODOS (Realtime Onboard Dependable Operating System)\*. The Onboard Data Processing course introduces students to the theory and application of the operating system. Students get to know the operating system as a tool for handling complexity. RODOS was developed at the German Aerospace Center (DLR) under the initiative and supervision of Sergio Montenegro. It is a very compact real-time operating system which, among other things, offers preemptive multitasking and middleware for communication. It has been written in C++ and allows object-oriented application development.

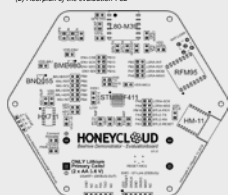
With regards to the cost-sensitive target application, a development hardware was created for HONEYCLOUD, based on commercially available off-the-shelf components, which was excellently suited as a cost-effective teaching platform. This platform consists of a relatively complex evaluation board and various wireless sensor systems. The architecture is designed in such a way that all variants of these wireless sensor systems represent a subset of the evaluation board. This means that any pin assignments or lines to subsystems, if any, are identical for all circuit variations. This allows a code or thread to be reused as a building block without the need for adjustments.

## Development hardware for research and teaching IoT systems in beekeeping



(1) Building blocks of the HONEYCLOUD embedded system

(2) Floorplan of the evaluation-PCB



(6) Fully functional beehive electronics demonstrator equipped with HONEYCLOUD sensor suite (Type D: IMU/Environmental /Scales-ADC/BTLE)



(7) Fully assembled evaluation-PCB



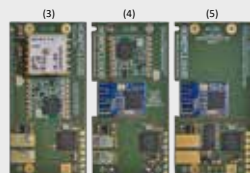
(8) Interior view of a weigh beam. A beehive is placed on two parallel weighing beams.



(9) Empty Langrath honey frames equipped with 3D printed sensor housings

## Results

Currently, there is an experimental IT infrastructure that can be built from cost-effective components. All typical IoT system components are covered. The platform consists of various wireless sensor nodes (edge devices) whose architecture is a subset of the on-site base station (data aggregator). This base station acts as a gateway and sends the data to a data center, which consists of a network of Linux-based single board computers (Raspberry Pi). The MongoDB distributed database and the Hadoop framework analysis run on this network. Future project work should enable an automated analysis of the data collected in the field test with the participation of beekeepers.



Variations of the HONEYCLOUD sensor nodes. From left to right: (3) Type C: LoRa/GPS-Beacon, (4) Type B: LoRa/Bluetooth LE Bridge, (5) Type A: Scales-ADC/Bluetooth LE

## Conclusions

The project is still in progress, but it has already shown that some aspects of the project are advantageous for use in teaching. Aerospace is always interdisciplinary. The study path of aerospace informatics merely determines that the topic is approached from the natural sciences (mathematics and computer science) and not, as is usually the case, via mechanical engineering. Students thus come from the field of construction, for example. In this way, student's works are prepared using parametric-shaped parts created with a geometry that is controlled by the properties of circuit diagrams. The interdisciplinary approach with the combination of aerospace and agriculture was originally planned as a strategy to increase innovation in research. It has been shown that this interdisciplinarity harmonizes very well with the interdisciplinary nature of the course.

## Further information

The University of Würzburg is seeking partners for exchange in research and teaching in the field of precision farming as well as for future international cooperation projects in the field of precision farming. We are looking forward to your feedback.

Contact: alexander.hilgarth@uni-wuerzburg.de

## Acknowledgments

Our most sincere gratitude and appreciation to Prof. Jürgen Tautz, who made this project possible. This project is funded by the and Education (BMBF) and is supervised by Federal Ministry of Education and Research Project Management Jülich (PU).



## IMPLEMENTING MACHINE LEARNING ALGORITHMS TO PREDICT DONOR STATUS: PRELIMINARY WORK WITH DATA FROM AN INSTITUTION OF HIGHER LEARNING

Cecilia Coulter / Paula Baingana / Pascaline Mukakamari

Identifying potential donors allows institutions of higher learning to conduct more effective fundraising campaigns. Machine learning classification algorithms can be useful in building models to predict donor status. However, when data contains imbalanced classes, like the data we used for this project, models tend to over-index the majority class, which was non-donors in this case. These results have significant implications for institutions in that they may not pursue entities that may, in fact, become donors. In order to improve the usefulness of our model, we used a resampling technique called random undersampling (RUS) to balance the data and also the area under the receiver operating characteristic curve (AUC-ROC) metric to evaluate the performance. Our final model improved its predictive power from 67% to 76%. Institutions of higher learning can use this machine learning model to more efficiently target the pool of potential donors, saving money and time. Future research will focus on improving the predictive accuracy of our model by exploring other data manipulation techniques that minimize the effect of imbalanced data, changing thresholds for classification algorithms, and using genetic programming and feature engineering.

### Implementación de algoritmos de aprendizaje automático para predecir el estado del donante

La identificación de posibles donantes permite a las instituciones de educación superior realizar campañas de recaudación de fondos más efectivas. Los algoritmos de clasificación de aprendizaje automático pueden ser útiles en la construcción de modelos para predecir el estado del donante. Sin embargo, cuando los datos contienen clases desequilibradas, como los datos que utilizamos para este proyecto, los modelos tienden a sobreindicar la clase mayoritaria, que en este caso eran los no donantes. Estos resultados tienen implicaciones significativas para las instituciones, ya que pueden no perseguir entidades que, de hecho, pueden convertirse en donantes. Para mejorar la utilidad de nuestro modelo, utilizamos una técnica de remuestreo llamada Random Under Sampling (RUS) para equilibrar los datos y utilizamos la métrica del área bajo la curva (AUC-ROC) para evaluar el rendimiento. Nuestro modelo final mejoró su poder predictivo del 67 % al 76 %. Las instituciones de educación superior pueden usar este modelo de aprendizaje automático para apuntar de manera más eficiente al grupo de donantes potenciales, ahorrando dinero y tiempo. La investigación futura se centrará en mejorar la precisión predictiva de nuestro modelo mediante la exploración de otras técnicas de manipulación de datos que minimicen el efecto de los datos desequilibrados, los umbrales cambiantes para los algoritmos de clasificación y el uso de la programación genética, así como la ingeniería de características.

# Implementing Machine Learning Algorithms to Predict Donor Status: Preliminary Work with Data from an Institution of Higher Learning



Cecilia Coulter, Paula Baingana, Pascaline Mukakamari  
mccoulter@stthomas.edu, ptbaingana@stthomas.edu, pascaline.mukakamari@stthomas.edu  
Graduate Programs in Software, University of St Thomas, Minnesota, USA

## Abstract

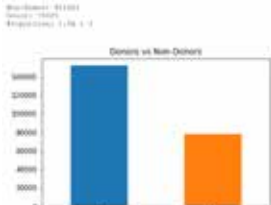
Identifying potential donors allows institutions of higher learning to conduct more effective fund-raising campaigns. Machine learning classification algorithms can be useful in building models to predict donor status. However, when data contains imbalanced classes, like the data we used for this project, models tend to over-index the majority class, which was the non-donors in this case. These results have significant implications for institutions in that they may not pursue entities that may, in fact, become donors. In order to improve the usefulness of our model, we used a resampling technique called random undersampling (RUS) to balance the data and also the area under the receiver operating characteristic curve (AUC-ROC) metric to evaluate the performance. Our final model improved its predictive power from 67% to 76%. Institutions of higher learning can use this machine learning model to more efficiently target the pool of potential donors, saving money and time. Future research will focus on improving the predictive accuracy of our model by exploring other data manipulation techniques that minimize the effect of imbalanced data, changing thresholds for classification algorithms, and using genetic programming and feature engineering.

## Introduction

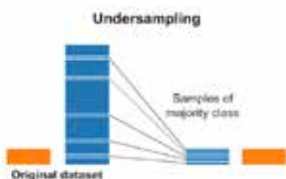
- **Problem:** What techniques can we employ with imbalanced data to develop a machine learning model that more accurately identifies potential donors?
- **Rationale:** Correcting for data imbalance before running classification algorithms should yield unbiased models that will be able to more effectively predict donor status.
- **Background:** Institutions of higher learning have plenty of data that would benefit from machine learning models that predict donor status. However, these data are sparse and imbalanced: two features that have rendered most models developed so far ineffective at predicting donor status.

## Materials and Methods

- **Data:** We dealt with over 200,000 records and 76 features. The majority class was Non-donors.



- **Resampling techniques:** We used random undersampling (RUS) (pictured) and synthetic minority oversampling technique (SMOTE).



- **Classification algorithms:** We ran logistic regression, decision tree, random forest, and naive-Bayes.
- **Dimensionality reduction:** We used principal component analysis (PCA) and linear discriminant analysis (LDA).
- **Evaluation metrics:** We explored confusion matrices, recall/precision/F1 tables, and AUC-ROC (receiver operating characteristic).
- **Language:** Python.

## Results

- **Preliminary results - model comparison.** Classification algorithms with raw data and 10-fold cross validation only are shown. Best model highlighted in orange.

Algorithm	Accuracy with k-fold
Logistic regression	74.7%
Decision tree	74.1%
Random forest	74.4%
Naive-Bayes	72.2%

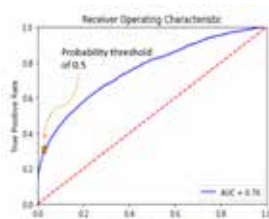
- **Logistic regression confusion matrix, no RUS.** Note the high number of false positives (in orange). With this model, most entities are predicted to be 'non-donors'.

	Predicted non-donor	Predicted donor
Actual non-donor	42927	2515
Actual donor	14997	8543

- **Logistic regression confusion matrix with RUS.** Note the improvement in false positives (in orange); fewer donors labeled as 'non-donors'. Accuracy falls to 68.77% but, at this point, we realize this metric is misleading (see next bullet).

	Predicted non-donor	Predicted donor
Actual non-donor	35463	9979
Actual donor	9349	14191

- **Evaluation - AUC-ROC for out best Model - logistic regression with RUS, our best model;** model evaluation improves to 76%.



## Conclusions

- **Best model:** logistic regression, 10-fold cross validation, RUS.
- **Best Evaluation metric:** AUC-ROC.
- **Typical donor profile:** Entities with strong ties to the institution, namely, having attended the institution, being a trustee, residing in the USA and planning, giving and attending events.



## Further Research

- **New algorithms:** Test new algorithms specifically designed to deal with imbalanced data.
- **New thresholds:** Change the threshold parameter to improve model's predictive power.
- **Data augmentation:** Merge with other databases, use APIs and harvest more information from the existing roster.

## Acknowledgements

We would like to thank Professor Michael Dorin, Clinical Professor, Graduate Programs in Software, UST, MN, for his guidance and encouragement, Dr. Manjeet Rege, Ph.D., Associate Professor, Graduate Programs in Software, UST, MN, who inspired us to undertake this project as part of a Machine Learning class and UST's Development Office for providing the data.

## References

- [1] Rafael Alencar. Resampling strategies for imbalanced datasets. <https://www.kaggle.com/raifjaa/resampling-strategies-for-imbalanced-datasets>, 2017.
- [2] Jason Brownlee. How and when to use roc curves and precision-recall curves for classification in python. <https://machinelearningmastery.com/roc-curves-and-precision-recall...>, 2018.
- [3] Renuka Joshi. Accuracy, precision, recall, f1 score: Interpretation of performance measures. [https://blog.exsilio.com/all/accuracy-precision-recall...](https://blog.exsilio.com/all/accuracy-precision-recall.../), 2016.

## OPEN-SOURCE SOFTWARE & PERSONAL MEDICAL DEVICES: INTERPRETING RISK THROUGH AN EVALUATION OF SOFTWARE TESTING

Michael Dorin / Heather Mortensen / Sergio Montenegro

The availability of powerful low-cost hardware and advanced software tools has made open-source medical devices possible. Deciding to use an open-source medical device may require acceptance of some risk. Fully comprehending the risk level is essential since failure of the software or the medical device is dangerous. As many medical applications contain complicated codes, an excellent method for understanding software readiness is to evaluate how much testing has been completed on the software. As a case study, this project evaluates the level of testing performed on software components of the Loop Artificial Pancreas system.

Classic methods for evaluating complicated source codes are used to demonstrate how much testing is needed in a project. Our analysis shows that the Loop Artificial Pancreas system (master branch) has been thoroughly tested with most of the faults likely discovered. By using classic software engineering metrics and techniques, it is possible to gauge how completely an open-source medical product has been tested and make an educated decision about the risk associated with using it.

### Software de código abierto y dispositivos médicos personales: interpretación del riesgo a través de una evolución de las pruebas de *software*

La disponibilidad de *hardware* de bajo costo y herramientas de *software* avanzadas ha hecho posible los dispositivos médicos de código abierto. La decisión de utilizar un dispositivo médico de código abierto puede requerir la aceptación de algún riesgo. Comprender completamente el nivel de riesgo es esencial ya que la falla del *software* o del dispositivo médico es peligrosa. Como muchas aplicaciones médicas contienen códigos complicados, un método excelente para comprender la preparación del *software* es evaluar la cantidad de pruebas que se han completado en el *software*. Como estudio de caso, este proyecto evalúa el nivel de pruebas realizadas en componentes de *software* del páncreas artificial de bucle.

Los métodos clásicos para evaluar el código fuente complicado se utilizan para demostrar cuántas pruebas se necesitan en un proyecto. Nuestro análisis muestra que el sistema de páncreas artificial de bucle (rama maestra) se ha probado exhaustivamente con la mayoría de las fallas probablemente descubiertas. Mediante el uso de técnicas y métricas de ingeniería de *software* clásico, es posible medir cuán completamente se ha probado un producto médico de código abierto y tomar una decisión informada sobre el riesgo asociado con su uso.



# Open Source Software & Personal Medical Devices: Interpreting risk through an evaluation of software testing



Michael Dorin, Heather Mortensen, Sergio Montenegro

mike.dorin@stthomas.edu, mort0048@stthomas.edu, sergio.montenegro@uni-wuerzburg.de



## ABSTRACT

The availability of powerful low-cost hardware and advanced software tools has made open-source medical devices possible. Deciding to use an open-source medical device may require acceptance of some risk. Fully comprehending the risk level is essential since failure of the software or the medical device is dangerous. As many medical applications contain complicated codes, an excellent method for understanding software readiness is to evaluate how much testing has been completed on the software. As a case study, this project evaluates the level of testing performed on software components of the Loop Artificial Pancreas system. Classic methods for evaluating complicated source codes are used to demonstrate how much testing is needed in a project. Our analysis shows that the Loop Artificial Pancreas system (master branch) has been thoroughly tested with most of the faults likely discovered. By using classic software engineering metrics and techniques, it is possible to gauge how completely an open-source medical product has been tested and make an educated decision about the risk associated with using it.

## CASE STUDY

### LOOP ARTIFICIAL PANCREAS

Loop was funded and built primarily by diabetic patients and caregivers. It was the first closed-loop system available in the US and preceded the release of a commercial system. Legal liability was limited by the fact that individual patients built the system themselves atop existing commercial medical devices.

Loop replaces an insulin pump controller and a continuous glucose monitor (CGM) receiver with an iPhone. A CGM sensor transmits blood glucose measurements via Bluetooth low energy (BLE) to the iPhone. The iPhone performs analytics and predictive statistics. Data is output to the user, alongside alerts for impending low blood sugar levels. Loop can utilize one of four different insulin models to perform the analysis. Loop incorporates data manually collected by the user inside Apple's Health app. In the event of Loop failure, insulin delivery reverts to default delivery rates in the pump. The iPhone sends BLE to the RileyLink, a custom-built piece of hardware that bridges communications between devices that use BLE (iPhone) and devices that use radio frequency (RF) (insulin pumps). RileyLink sends an RF command (916 MHz) to an insulin pump, where physical delivery of the medication to the user takes place through a catheter. [5] [6] [9]



## METHODOLOGY

1. Estimate potential bugs using Halstead method 1  
 $Bugs = (\text{Halstead effort}^{2/3})/3000$ . [2]  
 $Bugs = \text{Halstead volume}/3000$  [2]
2. Estimate potential bugs using Halstead method 2  
 $Bugs = \text{Halstead volume}/3000$  [2]
3. Estimate potential bugs using McConnell's method  
 A delivered bug every 1.5 to 50 lines of code. [4]
4. Akiyama's method  
 $Bugs = 4.86 + 0.018 X$  (lines of code) [1]
5. Estimate reported bugs reviewing online reporting.
6. Estimate reported bug count by scanning git log.
7. Compare estimated bugs to reported bugs.

## IMPORTANT ONGOING WORK

Loop is a mature project, so it was possible to evaluate it at a very tested state. However, new open-source medical and aerospace projects are released on a regular basis. More work is required to design a practical procedure using commonly available and affordable tools and techniques for evaluating risk on mission critical projects. The following aspects are being explored for the future procedure.

**McCabe's cyclomatic complexity** gives an indication of how many independent paths exist in a module by showing how many unit tests are required. Unit tests should be reviewed and execution reports analyzed [3].

**Static analysis** is a procedure for finding program faults using tools which examine the code without executing it.

**Human Complexity Analysis** measures the difficulties a person may encounter when reviewing the source code. The coding style is an important part of this.

**Weibull Analysis** will make a prediction based on bug discovery rates when about 2/3 of bugs are found. Weibull will be especially useful on new projects. [7]

## REFERENCES

1. Akiyama, F., An example of software debugging. IFIP Congress. (1971)
2. Halstead, Elements of software science. Vol. 2. New York Elsevier, 1977.
3. McCabe, "A complexity measure." IEEE Transactions on software Engineering 4 (1976)
4. McConnell, "Code Complete" Pearson Education (2004).
5. DiSimone, K., "My Artificial Pancreas", <https://myartificialpancreas.com/>, 2017
6. LoopDocs, <https://loopkit.github.io/loopdocs/#welcome-to-loop>, Retrieved July 2019.
7. Simmons, Erik, "Software Defect Arrival Modeling Using the Weibull Distribution." Pacific Northwest Software Quality Conference.
8. Fitzhery, pikaboy.com. Pictures retrieved July 2019
9. M1 pump frequency: <https://loopkit.github.io/loopdocs/faq/rileylink-faqs/>

Project	Online Reported Bugs	Bugs Keyword in Log
Amplitude-iOS	70	176
CGMBLEKit	28	52
G4ShareSpy	0	4
LoopKit	25	188
SwiftCharts	311	169
dexcom-share-client-swift	3	5
Rileylink iOs	67	347
Loop	458	415
<b>Total</b>	<b>972</b>	<b>1356</b>

Estimated Bugs	Technique	Online Reported	Keywords Reported
1284.156	Akiyama's Method	75.6 %	100 %+
1066.08	McConnell's Method	91.1 %	100 %+
936.16	Halstead 1	96.3 %	100 %+
1388.04	Halstead 2	70 %	97.6 %



## DISCUSSION AND CONCLUSION

The number of bugs found in a project is a good representation of how much progress has been made in the test workflow. With open-source projects, we found that not all bugs are reported to the official bug tracking software. Often developers make a fix and check it in without officially logging the bug. Also many times users of a particular project report bugs that do not inspire a fix from the developers. In this research, a scan of the git log as well as a review of on-line reported bugs were used to estimate bug fixes.

The particular tools and metrics used in this project were selected and evaluated because of their availability and ease of understanding for those with only a basic background in software development. It is imperative that a person considering the use of one of these devices has a practical means of evaluating their risk, without the requirement of complicated or expensive tools. All of the tools used in this project are open source.

This project and case study evaluated the state of software testing for a particular project. Software testing alone is essential, but it does not address all the areas of concern. Further research is necessary to build a complete set of practical tools and procedures. This project and case study have shown one important and practical method for evaluating open-source mission-critical systems.

## ACKNOWLEDGEMENTS

We wish to especially thank Kathryn DiSimone and Nate Racklyeft for their assistance with information and images, as well as necessary permissions.

## EMOTION TWENTY QUESTIONS IN CHINESE

Shanshan Kong / Abe Kazemzadeh

Our study introduces the emotion twenty questions (EMO20Q) game, an experiment into the cognition and expression of emotions in ordinary people who speak Chinese. The preliminary results show that such a game is felicitous and that the questions generated to describe emotions have commonalities with earlier studies conducted in English.

### Veinte preguntas sobre emociones en chino

Nuestro estudio presenta el juego Emotion Twenty Questions (EMO20Q), un experimento sobre el conocimiento y las expresiones de la emoción en personas comunes que hablan chino. Los resultados preliminares muestran que el juego es acertivo y que las preguntas generadas para describir las emociones tienen puntos en común con estudios anteriores de habla inglesa.

# Emotion Twenty Questions in Chinese

--A Preliminary Experiment for Dialogic Verbal Emotional Intelligence

Shanshan Kong, skong@stthomas.edu

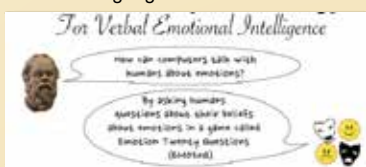
Abe Kazemzadeh, abe.kazemzadeh@stthomas.edu

**Abstract:**

Our study introduces the emotion twenty questions (EMO20Q) game, an experiment into the cognition and expression of emotions in ordinary people who speak Chinese. The preliminary results show that such a game is felicitous and that the questions generated to describe emotions have commonalities with earlier studies conducted in English.

**Introduction:**

EMO20Q is a dialog-based game that has been demonstrated to be helpful in studying emotion words [1]. It is similar to an ordinary twenty questions game, but focused on emotion words. The current EMO20Q game is restricted to support only English emotion words, which prevents its usage to study emotion words in other languages.



**Materials and methods:**

We made enhancements to the experiment designs to adapt to Chinese language characteristics. In each experiment, two players interacted using Chinese natural language, and each player had to pick two easy or difficult words to be guessed. We collected data from a human-human dialogue implemented in the mobile chat application Wechat. In the experiment instructions, it was emphasized that players could use synonyms, and hints were allowed.

**Literature cited:**

- [1] A. Kazemzadeh, P. G. Georgiou, S. Lee, and S. Narayanan.2011. "Emotion twenty questions: Toward a crowd-sourced theory of emotions". In: *Proceedings of ACL'11*.
- [2] Von Ahn L., Dabbish L. , "Designing games with a purpose", 2008, *Communication of the ACM*, col.51, no.8, pp.58-67.

**Acknowledgements:**

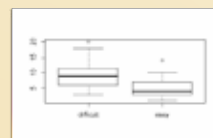
The authors would like to thank the EMO20Q players and the help of Professor Dorin Michael.



**Results:**

1. We collected a total of 44 games from 22 players. The number of questions in each game ranged from 1-18, and the mean number of questions was 7.2. The accuracy was 93.3%, including five successful synonym outcomes.

2. The number of easy and difficult words to be guessed was significantly different at  $p < .05$  (t-test, 3.369\*\*).



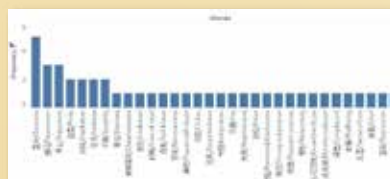
3. The most frequently mentioned emotion words were “高兴” (gladness)、“开心” (happiness) and “郁闷” (depression). Furthermore, four-ideogram emotion words such as “哭笑不得”(annoyed or embarrassed, possibly vicariously, to the point of laughing and crying) are unique to Chinese.

Table. Examples of question categories and question standardization

Standardized Question [Question Categories]	Examples
e==满足 [identity]	是满足吗？/Is it satisfied?
e.valence==positive [attribute]	心情好，对吧？/Is it a good emotion? 是一个积极的问题吗？/Is it positive?
cause(failure,e) [causal]	是因为事件失败而产生的吗？ /Is it because something failed?
e.frequency==regularly [attribute]	是经常会出现的情绪吗？/Is it easy to come by? 是生活中经常见到的情绪吗？/Is it a common emotion in life?
situation(e, stressState) [situational]	这种情绪是在应激状态下发生的吗？ /Does it happen in a stressful situation?
word.length==2 [morphology feature]	是两个字的词吗？/Does it contain 2 words?

**Conclusions:**

1. With this experimental design, it is possible to successfully study Chinese emotion words.
2. There may not be one-to-one translations between English and Chinese emotion words.
3. Further research studies, such as a pilot study of human-computer interaction and a large Mechanical Turk study, are needed. Finally, the aim is to define a Chinese computational model using a sequential Bayes pushdown automaton algorithm.



Graduate Programs in Software School of Engineering UNIVERSITY OF St. Thomas

## INTEGRATING AND ACCESSING UNIVERSITY INFORMATION APIS USING NATURAL LANGUAGE PROCESSING TOOLS

Shantanu Hadap / Shubha Shubha

Attending or working at a university requires students and faculty to manage a large amount of information to be successful. Both school administration and academic challenges have requirements to be met. One solution to this problem is to provide a virtual assistant which gives the required information to students and faculty. In this work, a virtual assistant that can completely understand conversational English and provide any needed information has been created. Talking to a virtual assistant to get the information is more convenient than the conventional way of doing it. Getting information from the virtual assistant does not require typing or browsing or any type of human interventions, which makes it more time-efficient and accessible. This solution has been tried in a private environment with a small scope at the University of St. Thomas and the results have been encouraging. Sixty percent (60%) of the users believe the assistant is very useful and the remainder finds it moderately useful.

### Integración y acceso a API de información universitaria utilizando herramientas de procesamiento de lenguaje natural

Asistir o trabajar en una universidad requiere que los estudiantes y la facultad administren una gran cantidad de información para tener éxito. Tanto la administración escolar como los desafíos académicos tienen requisitos que cumplir. Una solución a este problema es proporcionar un asistente virtual para estudiantes y profesores que brinde la información requerida. En este trabajo se crea un asistente virtual que puede comprender completamente el inglés conversacional y proporcionar la información necesaria. Hablar con un asistente virtual para obtener la información es más conveniente que la forma convencional de hacerlo. Obtener información del asistente virtual no requiere tipear o navegar ni ningún tipo de intervención humana, lo que lo hace más eficiente y accesible. Esta solución se ha probado en un entorno privado con un pequeño alcance en la Universidad de St. Thomas y los resultados han sido alentadores. El 60 % de los usuarios cree que el asistente es muy útil y el resto lo encuentra moderadamente útil.

# Integrating and Accessing University Information APIs using Natural Language Processing tools



**Shantanu Hadap, Shubha Shubha**  
 hada6576@stthomas.edu, shub6690@stthomas.edu  
 University of St. Thomas

## Abstract

Attending or working at a university requires students and faculty to manage a large amount of information to be successful. Both school administration and academic challenges have requirements to be met. One solution to this problem is to provide a virtual assistant which gives the required information to students and faculty. In this work, a virtual assistant that can completely understand conversational English and provide any needed information has been created. In this work, a virtual assistant is created that can completely understand conversational English and provide any needed information. Talking to a virtual assistant to get the information is more convenient than the conventional way of doing it. Getting information from the virtual assistant does not require typing or browsing or any type of human interventions which makes it more time-efficient and accessible. The implication of this work is that students, faculty and staff will have access to a virtual assistant to help out and thus minimize the stress of a school year. This solution has been tried in a private environment with a small scope at the University of St. Thomas and the results have been encouraging. Sixty percent (60%) of the users believe the assistant is very useful and the remainder finds it moderately useful.

## Introduction

Is it possible to have a virtual assistant to students and faculty that will completely understand conversational English and can provide the relevant answer or information? 'Canvas Panda' is one such assistant which is developed specifically for the University of St. Thomas. This app runs on smart assistants like Google Assistant or on web and mobile chat-bot applications like Facebook messenger and Skype. The user needs to specify "Talk to Canvas Panda" for the Google Assistant to take them to the app. This assistant app then gets the input questions from the users. These questions are mapped with the related intent defined in the Google Dialogflow. The intents can be with parameters or without parameters. The intents are mapped to the related methods in the cloud functions and these methods make the REST API calls to different university APIs like library, student/faculty directory, computer labs, etc. It can also access secured API services such as Canvas APIs [1] and student account APIs. The received information is cleaned and passed back to the assistant. The assistant provides the required information in English sentences as output.

## Materials and methods

Day to day university queries need a time-efficient solution with minimum user interactions and easy accessibility. To develop this solution, an NLP tool called Google Dialogflow [2] is used. An agent/app with an invocation name 'Canvas Panda' has been created in Google Dialogflow actions [3]. This app maps user conversational queries to specific intent using training phrases. Common queries based on university environment have been listed and intents have been created for each of these queries. Inside intents, training phrases have been defined. The training phrases are the different possible ways in which a query can be asked by the user. Dialogflow ML can map the query to intent even if it is not structured properly or is grammatically incorrect. Dialogflow also keeps on training and improving this mapping. The training phrases can also contain parameters/slots. The intents are linked to the Google Cloud functions [4] which contain methods for each intent. The cloud function is written using Node.js [3] and has been deployed on the Google Cloud platform. The cloud function calls university specific open APIs such as library service API, directory service API, etc. Integrating the third-party API can be achieved by implementing authentication methods like OAuth. The result obtained by API is cleaned and sent back to Google Assistant which reads or displays the response to the end user.



Figure 1. Architecture design of Canvas Panda. The APIs integrated are shown in green and the possible integrations are shown in yellow.

## Results

The preliminary version of this app has been tested in a private environment and a survey has been done with a small group of five people. The survey result shows that 60% of the users believe the assistant is very useful/accessible and the remainder finds it moderately useful/accessible. All the users find the app to be time-efficient. Users were asked to vote for their preferred choice of device for this application. The response was distributed as follows: 29% voted for smart-speaker, 29% for mobile, 18% for smartwatch, 12% for website and 12% for web messenger.



Figure 2. Survey results, text output using Google Assistant on mobile phone.

## Conclusion

Our main aim was to develop a time-efficient and easily accessible solution for day-to-day university queries of students and faculties. The project was initiated with a narrow scope and limited API integrations. The response of the people participating in the survey reflects that the draft version of the app build using Google Dialogflow satisfied our focused aspects of efficiency and accessibility.

The scope of this project could be further increased by including more actions/queries such as student account details, library systems, campus news and announcements, lab/study room bookings and availability, etc. Integration of different third-party university applications like Canvas, Outlook, etc. could be a future addition to this application.

## References

- [1] Instructure Inc. Canvas lms – rest api and extensions documentation.
- [2] Google Inc. Build an agent from scratch using best practices.
- [3] Google Inc. Build fulfillment with the Actions on Google Node.js client library.
- [4] Google Inc. Google Cloud Functions documentation.

## Acknowledgements

We would like to thank Michael Dorin and Abe Kazemzadeh for their guidance; and Brett Coup, Eric Tornoe and Benjamin Durrant for providing technical resources required for the project.

## P'AWAQ YUPANA – NEOÁBACO DE LÓGICA HÍBRIDA

Carlos Saldívar (Dhavit Prem)

El P'awaq Yupana es una herramienta lúdico-pedagógica desarrollada para facilitar la abstracción de la operación en base binaria en el uso de la yupana inca mediante el método Tawa Pukllay. Sirve, además, para potenciar el pensamiento computacional, permitiendo al usuario encontrar soluciones aritméticas mediante una lógica híbrida de tres tipos de razonamiento aritmético: el andino (reconocimiento de patrones), el indo-arábigo (cálculo mental numérico) y el chino-japonés (manejo de cuentas en herramientas de conteo “ábaco” como apoyo al cálculo mental).

### P'awaq Yupana: A New Hybrid Logic Abacus

P'awaq Yupana is a playful-pedagogical tool developed to facilitate the abstraction of a binary operation performed with the Inka Yupana through the Tawa Pukllay method. It is also useful to enhance computational thinking, enabling the user to find arithmetic solutions through the hybrid logic of three types of arithmetic reasoning: the Andean (pattern recognition), the Indo-Arabic (numerical mental calculation) and the Chinese-Japanese (handling of beads in a counting tool named “abacus” as a support for mental calculation).

# P'awaq Yupana - Neoábaco de lógica híbrida

Carlos Saldivar (Dhavit Prem)

dhavitprem@gmail.com,  
yupanki@yupanainka.com

Resumen: El P'awaq Yupana es una herramienta lúdico-pedagógica desarrollada para facilitar la abstracción de la operación en base binaria en el uso de la yupana inca mediante el método Tawa Pukllay. Sirve además para potenciar el pensamiento computacional, permitiendo al usuario encontrar soluciones aritméticas mediante una lógica híbrida de tres tipos de razonamiento aritmético: el andino (reconocimiento de patrones), el indo-arábigo (cálculo mental numérico) y el chino-japonés (manejo de cuentas en herramientas de conteo "ábaco" como apoyo al cálculo mental).

## Introducción

La yupana fue el calculador prehispánico utilizado por los incas, en combinación con los khipus: la yupana funcionaba para realizar las operaciones aritméticas, mientras que el khipu era el dispositivo para el almacenamiento de dichos datos. (Radicati, 1990)



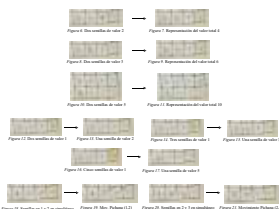
Figura 1. Curaca Condorchaua, por Guamán Poma de Ayala

Desde inicios de 1900 empezaron los primeros estudios sobre la escritura de números y operaciones aritméticas en la yupana (Pereyra, 1990). El método Tawa Pukllay (los cuatro juegos sagrados de los incas en quechua) (Prem, 2014), propone, en base a esos estudios, la decodificación de los cálculos mediante el reconocimiento de patrones y movimientos.

Así como el ábaco japonés, su uso en Occidente ha experimentado en los últimos tiempos gran desarrollo. Debido a la sencillez material, su naturaleza manipulativa ofrece la posibilidad de operar antes de trabajar con ellos de forma mental (Arteaga-Martínez, Sánchez, y Núñez del Río, 2016). Una de las ventajas de su uso es el gran refuerzo que ofrece desarrollando algo más que la exactitud y velocidad de cálculo aritmético, pero manteniendo la estructura de la yupana inca, adaptada para su operación mediante 1 bit por casillero.

## Materiales y métodos

El método Tawa Pukllay consiste en la resolución de operaciones aritméticas mediante el reconocimiento de patrones en movimientos y en paralelo, sin necesidad de cálculos mentales numéricos. Este se desarrolla en un tablero basado en el diseño de la yupana inca encontrada en las crónicas de Guamán Poma de Ayala, la cual es una matriz de cuatro columnas y cinco filas. Cada fila contiene la misma secuencia de números: 5, 3, 2, 1, un número por cada casilla. A su vez, cada fila representa una potencia del 10, para así representar las unidades, decenas, centenas, millares, decenas de millar y se puede continuar agregando filas de acuerdo a la cantidad a operar.



Partiendo de este método se desarrolló el P'awaq Yupana, en el cual se van ejecutando los movimientos conforme se van colocando los números en el tablero, de forma tal que cuando se ha terminado el último operando, se tiene ya la respuesta.

Debido a que en esta variante del Tawa Pukllay no se requiere más de una semilla por casilla, se creó un modelo físico distinto, basado en el diseño del sorobán japonés, pero manteniendo la estructura lógica de la yupana, añadiendo, además, a cada uno de ellos dos semiestados que facilitan el manejo de las operaciones.

El P'awaq Yupana se caracteriza porque integra tres tipos de lógica de manera alterna y en paralelo: manejo de ábaco, cálculo numérico y reconocimiento de patrones, lo cual enriquece el potencial de desarrollo de habilidades mentales que ofrece a sus practicantes.

## Resultados

Se creó una propuesta variante al sistema Tawa Pukllay en sumas y restas, esta vez aplicando el reconocimiento de patrones y movimientos en cada activación, de cada casilla, de cada operando, logrando así obtener la respuesta de manera inmediata una vez colocados todos los operandos.



Permitió, además, una adaptación a los procesos de pensamiento híbridos en el cálculo numérico, mediante un sistema sencillo de operar y lúdico.

Como también la abstracción de las operaciones a nivel binario en dicha estructura.

## Conclusiones

El P'awaq Yupana facilita el diseño lógico para la estructura de datos que representa y el consecuente desarrollo de tecnologías (hardware y software) basados en esta estructura y dinámica de datos.

Al ser el P'awaq Yupana un sistema que realiza cálculos aritméticos en paralelo y que ofrece dos semiestados en cada uno de los Ybits, ofrece grandes posibilidades para su estudio e implementación en el campo de la computación cuántica.

El P'awaq Yupana permite ser experimentado por personas jóvenes en edad escolar, aprendiendo nuevas técnicas de cálculo, reforzando las ya aprendidas y desarrollando abstracciones que faciliten el pensamiento computacional.

## Información adicional

© P'awaq Yupana 2018-2019 Todos los derechos reservados. P'awaq Yupana es una herramienta tecnológica educativa desarrollada por Dhavit Prem y la Asociación Yupanki.

## Referencias

Arteaga-Martínez, B., Sánchez, J., y Núñez del Río, M. (2016). Aprender con el ábaco japonés. Matemáticas, y... ¿algo más? XVI Congreso Nacional y VII Congreso Iberoamericano de Pedagogía: Democracia y Educación en el siglo XXI. Madrid.

Pereyra Sánchez, H. (1990). La Yupana, complemento operacional del quipu. En C. Mackey, Qhipu y Yupana. Lima: Concytec.

Poma de Ayala, F.G. (1615). *Primera nueva crónica y buen gobierno*.

Prem, D. (2014). *Yupana Inka - Decodificando la matemática Inka. Método Tawa Pukllay*. Asociación Yupanki.

Prem, D. (2018). P'awaq Yupana - El potente neo-awaku andino.

Radicati, C. (1990). Tableros de escaques en el Antiguo Perú. En C. Mackey, Qhipu y Yupana. Colección de escritos. Lima: Concytec.



## G SUITE PARA LA EDUCACIÓN EN LA UNIVERSIDAD PÚBLICA DE LA REGIÓN CUSCO DEL PERÚ

Hernán Nina-Hanco

Sobre la base de un estudio longitudinal, entre los años 2015 y 2018, el conjunto de aplicaciones de G Suite for Education tiende a incrementar su uso y se consolida como el principal soporte digital de la comunicación interna en la universidad pública de la Región Cusco, del Perú.

### G Suite for Education at Public Universities of the Cusco Region in Peru

Based on a longitudinal study conducted between 2015 and 2018, the G Suite for Education applications are increasingly being used and are consolidating their position as the main digital support for internal communication at public universities of the Cusco Region in Peru.



# G Suite for Education en la universidad pública de la Región Cusco del Perú

Hernán Nina-Hanco  
[hninaha@ulima.edu](mailto:hninaha@ulima.edu)

**Resumen.** Con base en un estudio longitudinal entre los años 2015 y 2018, el conjunto de aplicaciones de G Suite for Education incrementó su uso y se consolidó como el principal soporte digital de comunicación interna en la universidad pública de la Región Cusco del Perú.

## Introducción

Las universidades gestionan complejas infraestructuras de tecnologías de la información para proporcionar servicios digitales a un considerable número de usuarios. En este sentido, la computación en la nube se consolida como el principal proveedor tecnológico de *hardware* y *software* para las universidades. Por ejemplo, las universidades del Perú adoptaron soluciones en la nube de Google o Microsoft para externalizar el servicio de correo electrónico, y con ello muchos problemas operativos, de infraestructura, de soporte y de costos fueron absorbidos.

En particular, la Universidad Nacional de San Antonio Abad del Cusco (UNSAAC), atraída por la tendencia, el año 2013 migró a G Suite for Education (aplicaciones de Google). Sin embargo, los reportes de uso del primer año indicaron que la gran mayoría de los miembros de la comunidad universitaria no utilizaban la plataforma. Entonces, se desarrollaron una serie de estrategias de motivación para promover el uso de soportes digitales en la comunicación interna universitaria. Al final, se observó una mejora progresiva y sostenible del uso de G Suite for Education en la UNSAAC.

## Metodología e instrumentos

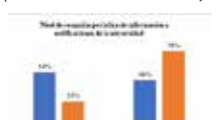
La investigación es de tipo aplicada con nivel de estudio descriptivo y explicativo. Asimismo, se utilizó el diseño cuasiexperimental con estrategia longitudinal con pre-test y post-test, con un solo grupo no aleatorio de una población de 20 000 usuarios. A continuación, se presenta el formato del diseño:

NA	G	O <sub>1</sub>	O <sub>2</sub>
----	---	----------------	----------------

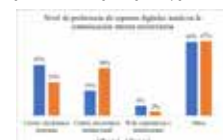
NA: muestreo no probabilístico o selección directa.  
 G: participantes en programa de motivación que son seleccionados de una población de 14 000 usuarios.  
 O<sub>1</sub>: observación en la etapa del pre-test antes de la aplicación del programa de motivación. Muestra seleccionada 368 integrantes. Ejecutado en el 2016.  
 O<sub>2</sub>: observación en la etapa del post-test después de la aplicación del programa de motivación. Muestra seleccionada 652 integrantes. Ejecutado en el 2018.  
 El instrumento utilizado es la encuesta con un cuestionario de preguntas cerradas con escalas de Likert. Además, para la validez y confiabilidad se emplearon pruebas de Alfa de Cronbach (0,73) y Aiken (0,016).

## Resultados

El nivel de recepción periódica de notificaciones de la universidad se incrementó en un 30 % según la comparación de resultados entre las etapas del pre-test y post-test realizadas durante los años de 2016 y 2018.



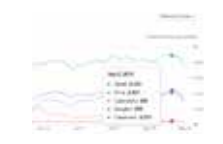
Con respecto al nivel de preferencia de soportes digitales, en la comunicación interna universitaria se destaca el uso del correo electrónico institucional, el cual se incrementó del 16 % al 30 %, según resultados comparados entre las etapas del pre-test y post-test.



Según los reportes de un período de seis meses, en el año 2015, se evidenció una constante de nivel bajo en el uso de aplicaciones de G Suite for Education.



Sin embargo, luego de aplicar una serie de programas de motivación, para utilizar soportes digitales en la comunicación interna, un reporte de un período de 6 meses, para el año 2018, mostró un incremento significativo y sostenible del uso de aplicaciones de G Suite for Education.



## Conclusiones

Según los resultados obtenidos y frente a las limitaciones en la infraestructura para el manejo de las tecnologías de la información en una universidad pública, G Suite for Education, en general, incrementó su uso, del 3 % al 18 %, de un total de 20 000 usuarios, entre los años 2015 y 2018.

Así mismo, G Suite for Education se consolidó como el principal soporte digital de comunicación interna en la Universidad Nacional de San Antonio Abad del Cusco. En primer lugar, G Suite for Education logró mejorar sustantivamente la frecuencia de notificaciones de información a usuarios en un 30 %. En segundo lugar, G Suite for Education potencia y simplifica, de manera sustantiva, el uso de soportes clásicos en la comunicación interna universitaria. La variedad de soportes clásicos como folletos, periódico, boletín o revista, buzón de sugerencias, entre otros, fueron absorbidos por las reuniones y tableros de anuncio que se gestionaron con herramientas digitales. En tercer lugar, G Suite for Education, acompañado de estrategias de motivación de comunicación interna, incrementó significativamente la utilización de soportes digitales en la comunicación interna universitaria, puesto que, el soporte digital con mayor crecimiento en preferencia es el correo electrónico institucional, el mismo que permitió el acceso a otras aplicaciones importantes como Google Classroom, Drive, Calendar, Meet. Finalmente, es importante puntualizar algunas limitaciones en la infraestructura informática y en las competencias digitales de los miembros de la comunidad universitaria UNSAAC. En ese sentido, con respecto a la infraestructura informática resulta que de cada 10 estudiantes 2 no poseen una computadora y la tercera parte del total de estudiantes accede al internet desde una cabina pública. Por otro lado, las competencias digitales adquiridas por docentes, administrativos y estudiantes son por su aprendizaje autodidacta o ayuda de compañeros. Sin embargo, estos hechos son frecuentes en muchas universidades públicas con bastante población estudiantil y limitada infraestructura informática.

## Recomendaciones

Es necesario crear un modelo de comunicación interna con soportes digitales basados en la integración de G Suite for Education y los sistemas de información desarrollados en la UNSAAC. Por ejemplo, extendiendo Google Classroom y Admin SDK se puede implementar un sistema de gestión de aprendizaje.

## Reconocimiento

Expreso mi reconocimiento a los miembros de la comunidad universitaria de la Universidad Nacional de San Antonio Abad del Cusco.

## Referencias

Alzate, A. G., y van Belle, J. P. (2016). The use of software as a service by students in higher education institutions: a systematic literature review. *IC2E '16 Proceedings of the 13th Annual International Conference on Electronic Commerce: e-Commerce in Smart Connected World*, 1-6. doi:10.1145/2914632.2914644

Al-Eman, M., y Iqbal Malik, S. (2016). The impact of Google apps at work: higher educational perspective. *International Journal of Innovative Mobile Technologies*, 10(3), 45-58. doi:10.3991/ijim.v10i03.4511

Baker, R., y Radovic, B. (2017). Cloud computing e-communication services in the university environment. *Information Systems Education Journal*, 15(1), 55-67.

Crum, G. E. (2016). Leveraging digital communication technology in higher education: exploring LRI's adoption of Google apps for education 2015. *Open Access Master's Thesis*. Pages 870. Recuperado de <http://digitalcommons.ac.edu/theses/970>

Mendon, S. H. V. (2015). Improving communication and building community with Google. *Proceedings of the 2015 ACM SIGCIS Annual Conference (SIGCIS '15)*. Association for Computing Machinery, New York, 85-90. doi:10.1145/2815862.2815874

Zajdman, M., y Saah, S. (2013). Software engineering frameworks for the cloud computing paradigm. *Reportado de https://link.springer.com/book/10.1007/978-1-4471-5011-2*





UNIVERSIDAD  
DE LIMA

Este libro recoge las conferencias magistrales y las ponencias presentadas durante el II Congreso Internacional de Ingeniería de Sistemas, denominado Innovando la Educación en Tecnología, organizado por la Carrera de Ingeniería de Sistemas de la Universidad de Lima, el cual tuvo lugar en la capital peruana los días 5 y 6 de septiembre del 2019.

Su temática principal giró alrededor del impacto de las tecnologías de la información en la educación. Las ponencias magistrales y los trabajos de investigación presentados ofrecieron innovadoras experiencias de aplicación y prometedoras líneas de trabajo futuro encaminadas a la transformación de las formas tradicionales de educación hacia esquemas dinámicos y soportados fuertemente por la tecnología.

La aplicación de paradigmas innovadores basados en técnicas de improvisación, gamificación, educación en ingeniería y enseñanza de la programación demostraron el potencial impacto que el binomio tecnología-educación puede generar en todos los niveles educativos y en diversos contextos de enseñanza-aprendizaje.