

Revista Killkana Técnica

Volumen 2, Número 1, Enero-Abril 2018

ISSN impreso: 2528-8024 ★ ISSN electrónico: 2588-0888



**UNIVERSIDAD
CATÓLICA DE CUENCA**
COMUNIDAD EDUCATIVA AL SERVICIO DEL PUEBLO

Cuenca, Junio de 2018

Revista Killkana Técnica

ISSN IMPRESO: 2528-8024

ISSN ELECTRÓNICO: 2588-0888

Dirección de Posgrado, Investigación
Vinculación con la Sociedad y Publicaciones
Universidad Católica de Cuenca

✉ Av. de Las Américas y Humbolt
Código Postal 010101, Cuenca - Ecuador

✉ killkana.editorial@ucacue.edu.ec

📞 Central telefónica:

+593 (07) 2-830-751

+593 (07) 2-824-365

+593 (07) 2-826-563

🌐 <http://www.ucacue.edu.ec>

Volumen 2, Número 1

Publicación trianual

Diseño, diagramación y maquetación en \LaTeX
Rodolfo Barbeito Rodríguez

English texts' reviewer
Tania Cecilia Bustamante Saavedra

Impresión: Editorial Universitaria Católica (EDÚNICA)

El sistema tipográfico empleado para componer la revista es \LaTeX , software libre utilizado para la comunicación y publicación de documentos científicos de alta calidad. Killkana Técnica emplea la clase `killkanaT.cls`, desarrollada especialmente para la revista y disponible para los autores en la página web <http://killkana.ucacue.edu.ec>



COMITÉ CIENTÍFICO INTERNO

Dr. Orlando Álvarez Llamaza, Phd.
Lic. Marilin Balmaseda Mederos, MSc.
Dr. Jaime Tinto Arandes, Phd.
Dr. Juan Carlos González, Phd.
Dr. Lorgio Aguilar Aguilar, Phd.
Dr. Santos Castillo Gaona, Phd.
Dr. Julio Conchado Martínez, Phd.
Dr. Mario Donoso Correa, Phd.
Dr. Ismael Minchala Ávila, Phd.
Dra. Milagro Rodríguez Andino, Phd.
Od. Esp. Lorena González Campoverde.
Dr. Edgar Curay Vanegas, Mgs.

COMITÉ CIENTÍFICO EXTERNO

Dr. Antonio Sacoto Salamea, Phd. / Universidad de la ciudad de Nueva York, Estados Unidos.
Dr. Carlos Parma, Phd. / Universidad del Aconcagua, Mendoza, Argentina.
Dr. Eber Orozco Guillén, Phd. / Universidad Politécnica de Sinaloa, México.
Dr. Adolf Sotoca, Phd. / Universidad Politécnica de Cataluña, España.
Dr. Gonzalo Salerno, Phd. / Universidad de Catamarca, Argentina.
Dr. Kai Tucci, Phd. / Universidad de Los Andes, Venezuela.
Dr. Carlos Echeverría, Phd. / Universidad de Los Andes, Venezuela.
Dr. Néstor Galán, Phd. / Universidad Politécnica de Sinaloa, México.
Dr. Julian Mora Aliseda, Phd. / Universidad de Extremadura, España.

DIRECTOR DE LA REVISTA

Ing. Rafael García Abad, MSc. / Director del Departamento de Posgrado, Investigación,
Vinculación con la Sociedad y Publicaciones

Prólogo

La unidad Académica de Tecnologías de la Información y la Comunicación de la Universidad Católica de Cuenca, en su proceso de formación de profesionales en las ramas de la computación, en su necesidad de contribuir a la socialización de aspectos relacionados con las carreras que desarrolla, en común acuerdo con la política institucional de contribuir al incremento del acervo científico, a través de la producción académica de sus colaboradores y técnicos relacionados, ejecuta el “*I Congreso Internacional Innovación Y Tecnología Informática en Las Organizaciones*”, el 27, 28 y 29 de junio de 2018.

Espacio, que tiene como intención principal aportar al desarrollo organizacional, a través de la divulgación de trabajos científicos sobre la informatización en las organizaciones, elaborados en base a políticas y estándares de revisión y edición, factores que garantizan la calidad de la producción. Los conjuntos de ponencias validadas se enmarcan en las temáticas del evento, que contemplan: sistema de información organizacionales; inteligencia de negocios; software organizacional; calculo computacional; internet de las cosas; seguridad informática; gobierno y gestión de tecnologías de información; entre otras. Obras que se sintetizan en el presente ebook, medio digital que busca trasladar fronteras como una fuente válida de conocimiento que apoya el progreso de la comunidad en su conjunto.

Paralela a la producción de la obra, se desarrolla el congreso, con ciclos de ponencias, talleres prácticos y charlas magistrales, dirigidas a profesionales, profesores y estudiantes con afinidad en la rama; entre los expositores destacan, profesionales e investigadores con una amplia trayectoria internacional y experiencia.

Que todo lo en esta obra se comparte sea conocimiento válido, que nos permita crecer profesionalmente, que sea parte de los instrumentos que apoyen el fortalecimiento y madurez del sector informático de la región y el país, en pos de apoyar la consecución de mejores condiciones de vida para el grupo humano, con el que cohabitamos.

Ing. Diego Marcelo Cordero Guzmán, PhD.
Decano de la Unidad Académica de
Tecnologías de la Información y Comunicación
Universidad Católica de Cuenca

Índice general

Prólogo	V
Propuesta de un sistema de aprendizaje de lengua de señas basado en las tecnologías de la información y comunicación <i>Milton Alfredo Campoverde Molina, Jacqueline Mishel Reina Alvarado, Antonio Gabriel Caranguí Delgado</i>	1
Evaluación de seguridad de la información en las páginas web pertenecientes a los municipios de la provincia del Cañar <i>Cristhian Flores Urgilés, Beatriz Zhinin Aguayza, Alexandra Segovia Cantos, Mercedes Mayancela Zhinin, Jessica Marlene García</i>	13
Revisión de algoritmos para la detección de valores atípicos <i>Cristina Mariuxi Flores Urgiles y Martin Sebastian Ortiz Amoroso</i>	19
Modelos de Referencia de Arquitectura Empresarial para la Industria de Educación Superior <i>Blanca Lucia Avila Correa</i>	27
Detección de filtraciones de privacidad: una visión del estado del arte <i>Diana Marcela Romero Córdova</i>	35
Instructivo para la preparación de manuscritos en la revista Killkana Técnica	41
Cómo usar la clase de documento L ^A T _E X ‘killkanaT.cls’ de la revista Killkana Técnica	47
Sistema y proceso de arbitraje de la revista Killkana Técnica	57

Propuesta de un sistema de aprendizaje de lengua de señas basado en las tecnologías de la información y comunicación

Proposal of a sign language learning system based on information and communication technologies

Milton Alfredo Campoverde Molina ^{1*}, Jacqueline Mishel Reina Alvarado ¹, Antonio Gabriel Carangui Delgado ¹
Unidad Académica de Tecnologías de la Información y la Comunicación (TIC)
Cuenca, Ecuador ¹

* mcampoverde@ucacue.edu.ec

DOI: https://doi.org/10.26871/killkana_tecnica.v2i1.284

Resumen

Esta investigación revela el desarrollo del software que con uso de dispositivos de reconocimiento de voz traduce a un lenguaje de señas; estableciendo como objeto de estudio a 10 estudiantes con discapacidad auditiva de la Unidad Educativa Especial Claudio Neira Garzón: Segundo de Básica (1 estudiante), Cuarto de Básica (5 estudiantes) y Quinto de Básica (4 estudiantes), a continuación, se realizó un análisis del porcentaje de discapacidad de cada estudiante de acuerdo al carnet del Consejo Nacional para la Igualdad de Discapacidades (CONADIS), luego se aplicó una encuesta a los estudiantes y se mantuvo una entrevista con la directora y subdirectora de la Unidad Educativa, determinando que sería de gran apoyo utilizar alternativas tecnológicas en el proceso de enseñanza-aprendizaje de la lengua de señas en los niños de 7 y 15 años con déficit auditivo parcial o total dentro de la asignatura de Lengua y Literatura en la lectoescritura, para el desarrollo del software se utilizó la metodología en cascada, luego de haber desarrollado el sistema se hicieron varias pruebas obteniendo muy buenos resultados en cuanto a su uso, incrementando el nivel de comprensión y entendimiento del lenguaje español y la lengua de señas en los niños y además hubo una excelente acogida por parte de los docentes y estudiantes que demostraron mucho interés por usar el software en el proceso de enseñanza-aprendizaje.

Palabras clave: Discapacidad auditiva, educación, lenguaje señas, software, TIC.

Abstract

This research reveals the development of software that translates into sign language with the use of speech recognition devices; establishing as object of study 10 students with hearing disability of the Claudio Neira Garzón Special Educational Unit: Basic Second (1 student), Basic Room (5 students) and Basic Fifth (4 students), followed by a analysis of the percentage of disability of each student according to the card of the National Council for the Equality of Disabilities (CONADIS), then a survey was applied to the students and an interview was held with the director and assistant director of the Educational Unit, determining what would be of great support to use technological alternatives in the teaching-learning process of sign language in children of 7 and 15 years with partial or total auditory deficit within the subject of Language and Literature in reading and writing, for the development of software He used the methodology in cascade, after having developed the system, several tests were made obtaining very good results. n in terms of its use, increasing the level of understanding and understanding of the Spanish language and sign language in children and in addition there was an excellent reception from teachers and students who showed great interest in using the software in the teaching process -learning.

Key words: Hearing impairment, education, sign language, software, TIC.

I. INTRODUCCIÓN

La Constitución de la República del Ecuador 2008 [1] en el artículo 26 considera que “la educación es un derecho de las personas a lo largo de su vida y un deber ineludible e inexcusable del estado. Del mismo modo, en el artículo 46, numeral 3 estipula la “atención preferente para la plena integración social de quienes tengan discapacidad.

El estado garantizará su incorporación en el sistema de educación regular y en la sociedad”. También, reconoce a las personas con discapacidad, los derechos a: (7.-) Una educación que desarrolle sus potencialidades y habilidades para su integración y participación en igualdad de condiciones. Se garantizará su educación dentro de la educación regular. Los establecimientos educativos cumplirán

normas de accesibilidad para personas con discapacidad e implementarán un sistema de becas que responda a las condiciones económicas de este grupo.

De la misma forma, en el Plan Nacional de Desarrollo 2017-2021. Toda una Vida [2] se plantea como primer objetivo “Garantizar una vida digna con iguales oportunidades para todas las personas”, estableciendo como una de sus Políticas numeral 1.4. “Garantizar el desarrollo infantil integral para estimular las capacidades de los niños y niñas, considerando los contextos territoriales, la interculturalidad, el género y las discapacidades” y como una de sus metas “dotar de ayudas técnicas a las personas con discapacidad a 2021”; asimismo, el CONADIS [3] en sus normas jurídicas en discapacidad Ecuador establece:

- “Acceso a una educación inclusiva y especializada de calidad a las personas con discapacidad que atienda sus necesidades específicas” (educación), así como el “Acceso al medio físico, Servicios de transporte, Tecnologías de la información y comunicación”.
- Dotar o repotenciar la infraestructura, el equipamiento, la conectividad y el uso de TIC, recursos educativos y mobiliarios de los establecimientos de educación pública, bajo estándares de calidad, adaptabilidad y accesibilidad, según corresponda. (4.1.c.-PNBV).
- Una educación que desarrolle sus potencialidades y habilidades para su integración y participación en igualdad de condiciones. Se garantiza su educación dentro de la educación regular. Los planteles regulares incorporarán trato diferenciado y los de atención especial la educación especializada. Los establecimientos educativos cumplirán normas de accesibilidad para personas con discapacidad e implementarán un sistema de becas que responda a las condiciones económicas de este grupo.

La OMS en su informe mundial sobre la discapacidad 2011 estima que más de mil millones de personas viven con algún tipo de discapacidad; o sea, alrededor del 15 % de la población mundial (según las estimaciones de la población mundial en 2010). Esta cifra es superior a las estimaciones previas de la Organización Mundial de la Salud, correspondientes a los años 1970, que eran de aproximadamente un 10 % [4].

Según, las estadísticas publicadas por el Consejo Nacional para la Igualdad de Discapacidades (CONADIS) [5] con información del Ministerio de Salud Pública de las personas registradas con discapacidad en el Ecuador, Azuay y Cuenca se puede ver la tendencia de crecimiento de las discapacidades: auditiva, física, intelectual, lenguaje, psicosocial y visual [6] [7] [8] en el tiempo en las Fig. 1, 2, 3 respectivamente:

Fig. 1. Tendencia de la discapacidad en Ecuador.

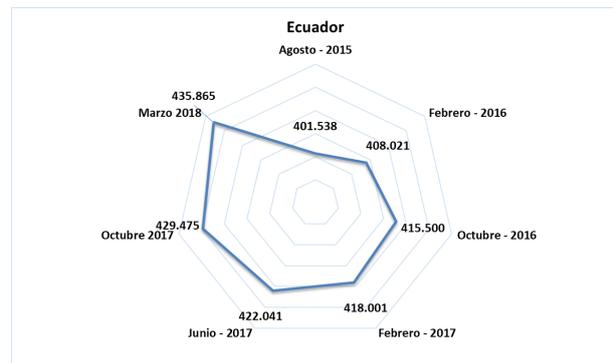


Fig. 2. Tendencia de la discapacidad en el Azuay.

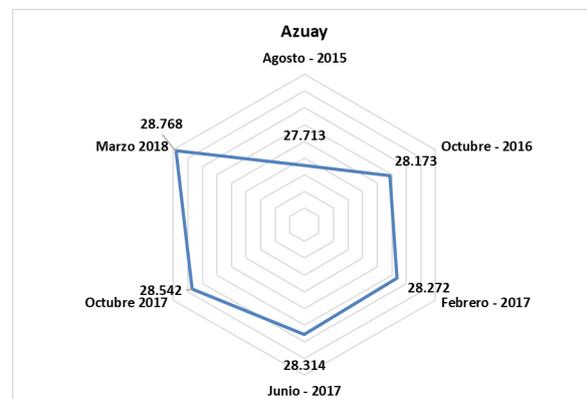
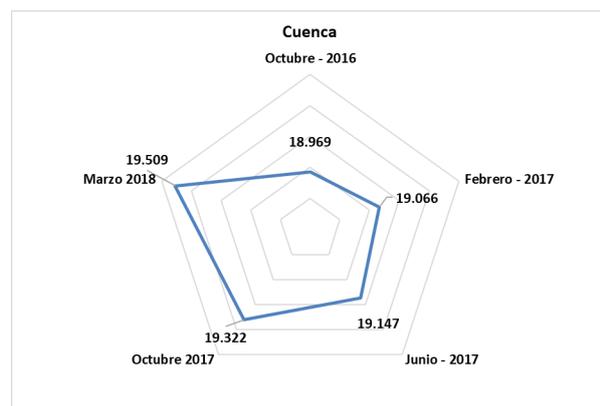


Fig. 3. Tendencia de la discapacidad en el Cuenca.



De la misma manera, en base a las estadísticas publicadas por el Consejo Nacional para la Igualdad de Discapacidades (CONADIS) [5] con información del Ministerio de Salud Pública de las personas registradas con discapacidad en el Ecuador, Azuay y Cuenca se puede ver la tendencia de crecimiento de la discapacidad auditiva [6] [7] [8] en el tiempo en las Fig. 4, 5, 6 respectivamente:

Fig. 4. Tendencia de la discapacidad auditiva en el Ecuador.

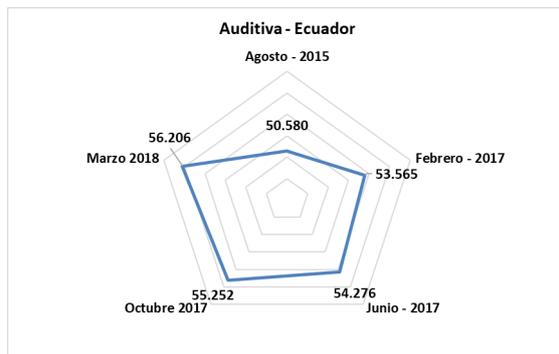


Fig. 5. Tendencia de la discapacidad visual en el Ecuador.

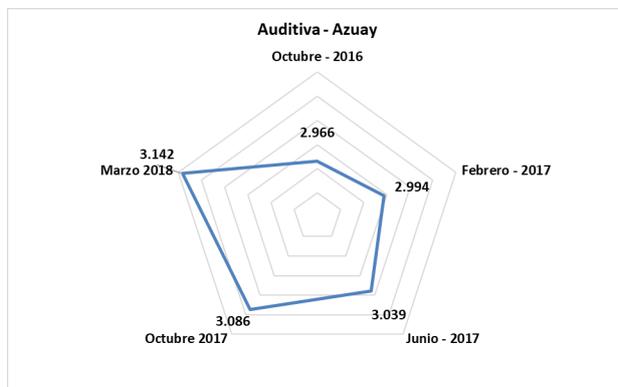
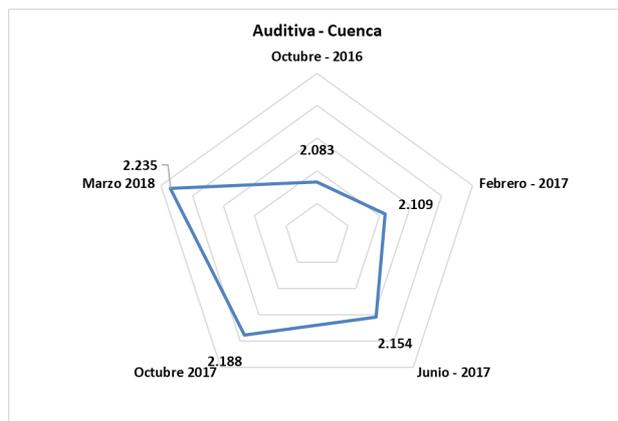


Fig. 6. Tendencia de la discapacidad auditiva Cuenca - Ecuador.



Haciendo, un análisis de los gráficos expuestos anteriormente se puede ver el crecimiento en número de personas con discapacidad en el tiempo en el Ecuador, Azuay y Cuenca; teniendo en cuenta, que las causas [9] de discapacidades en el Ecuador han sido originadas por: accidente de trabajo, accidente de tránsito, accidente deportivo, accidente doméstico, congénito/genético, desastres naturales/guerra, enfermedad adquirida, problemas de parto, trastornos nutricionales, violencia y otros (no se sabe), razón por lo cuál esta investigación busca alternativas

tecnológicas que fomenten un aporte significativo en la enseñanza-aprendizaje de niños con discapacidad auditiva. Contemplando, que en el ámbito educativo las TIC son consideradas valiosas herramientas de trabajo que pueden contribuir al desarrollo personal con un carácter individualizado e independiente pero integracionista y creativo. Por otra parte, su utilización en las distintas actividades económicas, sociales, culturales, entre otras, ha impuesto la necesidad de que la escuela asuma un rol activo en la preparación del futuro ciudadano para la utilización de dichas tecnologías, de manera que pueda utilizarlas para su desarrollo [10]. Si se quiere, lograr una educación inclusiva [11], es necesario determinar cómo deben ocupar ese espacio las personas sordas. En primer lugar hay que establecer que el problema más grande que enfrentan estos estudiantes es un problema de comunicación. Sin la posibilidad de comunicarse y acceder a la información que se maneja en las escuelas, es difícil, si no imposible acceder a “una educación con equidad”, si no hay un idioma común con el cual se pueda comunicar con su maestro o maestra.

Considerando todos los puntos tratados en los párrafos anteriores, esta investigación plantea como objetivo desarrollar un software que con uso de dispositivos de reconocimiento de voz traduzca a lenguaje de señas y sirva como herramienta de trabajo en la labor educativa de la Unidad Educativa Especial Claudio Neira Garzón en beneficio de sus niños con discapacidad auditiva.

II. MARCO TEÓRICO

A. Discapacidad Auditiva

La Ley Orgánica de Discapacidades del Ecuador [12] en el artículo 6 establece que "se considera persona con discapacidad a toda aquella que, como consecuencia de una o más deficiencias físicas, mentales, intelectuales o sensoriales, con independencia de la causa que la hubiera originado, ve restringida permanentemente su capacidad biológica, psicológica y asociativa para ejercer una o más actividades esenciales de la vida diaria"; también, en el artículo 33 se revela que la autoridad educativa nacional en el marco de su competencia, vigilará y supervisará, en coordinación con los gobiernos autónomos descentralizados, que las instituciones educativas escolarizadas y no escolarizadas, especial y de educación superior, públicas y privadas, cuenten con infraestructura, diseño universal, adaptaciones físicas, ayudas técnicas y tecnológicas para las personas con discapacidad, adaptación curricular, participación permanente de guías intérpretes, según la necesidad y otras medidas de apoyo personalizadas y efectivas que fomenten el desarrollo académico y social de las personas con discapacidad.

La discapacidad auditiva se define como la dificultad que presentan algunas personas para participar en actividades propias de la vida cotidiana, que surge como consecuencia de la interacción entre una dificultad específica para percibir a través de la audición los sonidos del ambiente y dependiendo del grado de pérdida auditiva, los sonidos

del lenguaje oral, y las barreras presentes en el contexto en el que se desenvuelve la persona. Tener una dificultad auditiva no implica tener restricciones para comunicarse. Una cosa es el proceso de comunicación y otra el medio y el canal empleados para hacerlo. Lo que una persona con dificultades auditivas puede tener son restricciones para la utilización del lenguaje oral como medio exclusivo de comunicación, sin embargo, las posibilidades de comunicar mediante otros medios están intactas. Por este motivo, las personas sordas desarrollan medios diferentes para comunicarse, como la lengua de señas [13].

B. Lengua de Señas

Hoy en día los niños con discapacidad auditiva [14] presentan grandes dificultades para comunicarse, ya que su principal lenguaje de comunicación es el lenguaje de señas, pero es necesario enseñar a los niños a comunicarse a través de la lectura y la escritura para incorporarse dentro de la sociedad.

La Ley Orgánica de Discapacidades del Ecuador [12] en el artículo 70 reconoce la lengua de señas ecuatoriana como lengua propia y medio de comunicación de las personas con discapacidad auditiva. Se incorporará progresivamente el servicio de intérpretes de la lengua de señas ecuatoriana en las instituciones públicas, así como la capacitación de las y los servidores públicos en la misma. La lengua de señas [15] del país es conocida actualmente como Lengua de Señas Ecuatoriana, LSE. En los textos escritos sobre esta lengua en inglés se la representa con las siglas LSEC.

El lenguaje de señas constituye el principal recurso para la comunicación dentro de la comunidad sorda, que la aprende de forma natural y la convierte en una lengua necesaria, útil y práctica [16].

C. Discapacidad Sensorial

Los alumnos y alumnas con discapacidad sensorial son aquellos que presentan una deficiencia en su audición o visión con resultado de una percepción y procesamiento de la información disfuncionales, dificultándoles aspectos cognitivos en la adquisición y manipulación del espacio, en el lenguaje, en la relación personal y de funcionamiento en la comunidad. Por ello, y para que su desarrollo personal evolucione en el sentido general o de la normalidad poblacional, precisarán de los apoyos y recursos que compensen sus dificultades y refuercen sus habilidades. Dada la importancia que tienen los canales auditivo y visual en la adquisición de la información, se pasará a revisar sus dos discapacidades [17].

D. Lectoescritura

A lo largo de la historia, la concepción y status social de la lectoescritura ha evolucionado pasando de ser un "conocimiento" de los más privilegiados, a un aprendizaje fundamental para la integración en la sociedad actual, en la cual, autonomía y eficacia son dos condiciones esenciales [18]. La escalera de ubicación propuesta por Grossi, indica

los conocimientos y habilidades en cada nivel, en los cuales ascender es sinónimo de aprendizaje y evolución en el proceso cognitivo [19], así:

- Nivel Presilábico 1: El estudiante expresa su conocimiento a través de dibujos y/o grafismos no convencionales.
- Nivel Presilábico 2: El estudiante emplea códigos similares con regularidad para expresar las mismas ideas, empieza a establecer códigos definidos.
- Nivel silábico: El estudiante tiene en cuenta el número de sílabas de la palabra y de acuerdo a ello coloca la cantidad de símbolos.
- Nivel Alfabético: El estudiante emplea códigos reconocidos y aceptados para expresarse pero comete adiciones u omisiones dentro de algunas palabras, en este proceso comienza a perfeccionar su nivel escritor.
- Nivel Alfabetizado: En este nivel los estudiantes escriben y reconocen los códigos establecidos para la comunicación en su lengua y es capaz de producir textos ya sea bajo orientación o por su propia iniciativa. Es en este nivel donde se determina que el estudiante ha logrado un aprendizaje y dominio del proceso lectoescritor.

El desarrollo de la lectoescritura implica los siguientes pasos en el proceso de la conciencia cognitiva: primero, pasar de la no-conciencia de la relación entre la escritura y el lenguaje hablado; a asociar lo escrito con el lenguaje oral; y al dominio de los signos escritos referidos directamente a objetos o entidades. Segundo, pasar del proceso de operaciones conscientes como la individualización de los fonemas, la representación de estos fonemas en letras, la síntesis de las letras en la palabra, la organización de las palabras; a la automatización de estas operaciones; y al dominio del texto escrito y del lenguaje escrito. El lenguaje escrito es una forma compleja de actividad analítica, en la cual la tarea fundamental es la toma de conciencia de la construcción lógica de la idea [20].

E. Las Tecnologías de la Información y la Comunicación en la educación

La educación se encuentra en un proceso de transformación de los procesos de enseñanza-aprendizaje para adaptarse a los cambios tecnológicos de la sociedad del siglo XXI, por la motivación de docentes e investigadores que han visto en las tecnologías un motor de cambio del sistema educativo y del quehacer de sus agentes [21].

Por tal razón, los currículos de la Educación General Básica (EGB) y el Bachillerato General Unificado (BGU) [22] que se implementan mediante Acuerdo Nro. MINEDUC-ME-2016-00020-A del 17 de febrero de 2016 en el Ecuador, tienen como prioridad el uso habitual de las tecnologías de la información y de la comunicación como instrumento facilitador para el desarrollo del currículo en cada una de las áreas. Para el efecto, de acuerdo a un análisis pedagógico y curricular, este proceso puede impulsarse desde los primeros años escolares y fortalecerse en la Básica Superior

y Bachillerato, con la finalidad de que los estudiantes desarrollen destrezas suficientes que le faciliten el uso de los programas básicos de ofimática, como herramientas tecnológicas y didácticas, para su aprendizaje, además de fomentar el conocimiento en las Ciencias de la Computación. Con esta finalidad, dependiendo de la disponibilidad de cada una de las instituciones, los docentes de todas las áreas tanto del nivel de Básica como del Bachillerato, deberán planificar y desarrollar sus clases de las diferentes asignaturas utilizando los laboratorios de informática, es necesario organizar un horario de uso de laboratorios, garantizando su máxima capacidad durante la jornada escolar. Además, de sus horas pedagógicas, el docente de informática deberá programar capacitaciones para los docentes de las demás especialidades que garanticen su actualización digital.

F. La discapacidad auditiva en el Ecuador

Según, los antecedentes de personas con discapacidad auditiva en la educación [15]:

- 1940: Enriqueta Santillán inicia actividades docentes con niños sordos en Quito.
- 1952: Se funda el Instituto Mariana de Jesús, por iniciativa privada. El instituto atendía niños con discapacidad visual y con discapacidad auditiva.
- 1962: se funda el Instituto Nacional de Audición y Lenguaje (INAL), en Quito, a partir de lo que fue una sección para niños con necesidades especiales en el Colegio Espejo.
- 1966: Varios jóvenes sordos comienzan a agruparse en Quito para organizar eventos deportivos.
- 1975: Se crea la Sociedad de Sordos Adultos Fray Luis Ponce de León, en Quito.
- 1976: Una sociedad similar a la quiteña se crea en Guayaquil.
- 1982: Se inicia el Proyecto “Mano a Mano” (estudio de la lengua de señas local) en la Sociedad Ponce de León.
- 1984: Se crea la asociación de sordos de Guayas.
- 1986: Se funda la Federación Nacional de Sordos de Ecuador.
- 1988: Aparece como libro una compilación de señas usadas por los sordos de Ecuador, resultado del Proyecto “Mano a Mano”.
- 2012: Se publican el Diccionario Oficial de Lengua de Señas Ecuatoriana y un estudio sociolingüístico sobre la LSE (SIL International).

G. Inclusión Educativa

La educación inclusiva es el proceso que ofrece a todos los niños y niñas, sin distinción de discapacidad, raza o cualquier otra diferencia, la oportunidad para continuar siendo miembro de la clase ordinaria y para aprender de sus compañeros, juntamente con ellos, dentro del aula [23].

La UNESCO define la educación inclusiva como el proceso de identificar y responder a la diversidad de las necesidades de todos los estudiantes a través de la mayor

participación en el aprendizaje, las culturas y las comunidades, y reduciendo la exclusión en la educación. Involucra cambios y modificaciones en contenidos, aproximaciones, estructuras y estrategias, con una visión común que incluye a todos los niño/as del rango de edad apropiado y la convicción de que es la responsabilidad del sistema regular, educar a todos los niño/as [24].

Hacer efectivo el derecho a la educación exige garantizar que todos los niños, niñas y jóvenes tengan, en primer lugar, acceso a la educación, pero no a cualquier educación sino a una de calidad con igualdad de oportunidades. Son, justamente, esos tres elementos los que definen la inclusión educativa o educación inclusiva. Podríamos decir entonces, que la inclusión educativa es consustancial al derecho a la educación o, en forma más definitoria, un requisito del derecho a la educación, toda vez que el pleno ejercicio de este derecho, implica la superación de toda forma de discriminación y exclusión educativa[25].

La inclusión educativa presenta una complejidad que puede ser comprendida de mejor manera si se tiene atención sobre el profesor como agente relevante y clave de este proceso. Puede constituirse en una barrera o en un agente facilitador de las prácticas inclusivas. Las actitudes del profesor acerca de la inclusión educativa, entendidas como el conjunto de percepciones, creencias, sentimientos y formas de actuar, impactan la disposición hacia la inclusión de personas con NEE. Una actitud positiva hacia prácticas inclusivas va a favorecer dicho proceso. En tanto una actitud negativa minimizará las oportunidades de aprendizaje y participación de estudiantes con algún tipo especial de necesidades educativas [26].

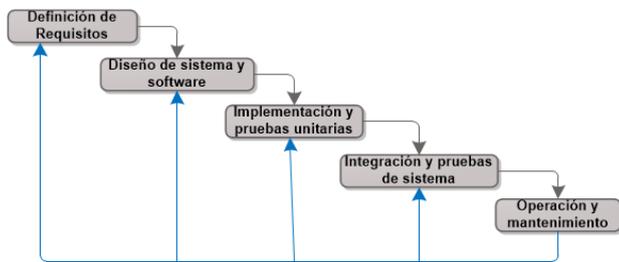
III. METODOLOGÍA

La metodología utilizada en el desarrollo del software de aprendizaje de la lengua de señas mediante reconocimiento de voz es la metodología en cascada que consta de las siguientes fases [27]:

- 1) **Definición de los requisitos:** los servicios, restricciones y objetivos son establecidos con los usuarios del sistema. Se busca hacer esta definición en detalle.
- 2) **Diseño de software:** se divide el material educativo computarizado en sistemas de software o hardware. Se establece la arquitectura total del material. Se identifican y describen las abstracciones y relaciones de los componentes del sistema.
- 3) **Implementación y pruebas unitarias:** construcción de los módulos y unidades del material educativo computarizado. Se realizan pruebas de cada unidad.
- 4) **Integración y pruebas del sistema:** se integran todas las unidades, se prueban en conjunto, se entrega y aplica el conjunto probado al cliente.
- 5) **Operación y mantenimiento:** generalmente es la fase más extensa. El material educativo computarizado es puesto en marcha y se realiza la corrección de errores descubiertos. Se realizan mejoras de implementación. Se identifican nuevos requisitos.

La interacción entre fases puede observarse en la Fig. 7. Cada fase tiene como resultado documentos que deben ser aprobados por el usuario:

Fig. 7. Modelo de desarrollo en cascada [27].



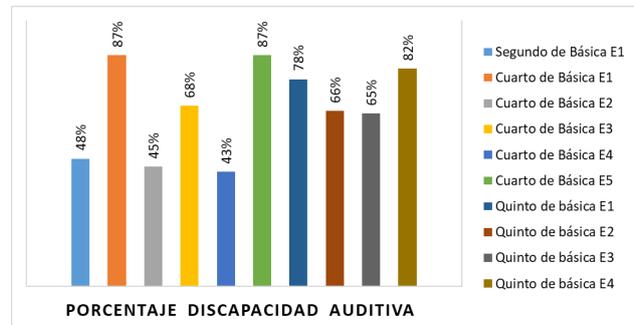
IV. RESULTADOS

A. Definición de los requisitos

Esta investigación forma parte del proyecto de investigación científica “La accesibilidad tecnológica como estrategia de inclusión educativa”, ejecutado por docentes investigadores y estudiantes de la carrera de Ingeniería de Sistemas de la Universidad Católica de Cuenca; la misma que se está desarrollando en la Unidad Educativa Especial Claudio Neira Garzón. Según su reseña histórica se puede ver que en el año de 1971 se crea con el nombre de Instituto Fiscal Especial de Invidentes y Sordos del Azuay, con Acuerdo Ministerial No.16, dictado por la Dirección Provincial de Educación, desde entonces han tomado en cuenta la necesidad social y humana de dar oportunidad de educación y cultura a las personas y niños con capacidades especiales físicas o intelectuales que les impidan desarrollarse y a su vez ser útiles en la sociedad. En el año 2014 se dio el cambio de denominación del establecimiento, según la normativa legal vigente como Unidad Educativa Especial Claudio Neira Garzón, nombre que representa al fundador Dr. Claudio Neira Garzón, profesor de invidentes desde enero de 1972 hasta diciembre de 1980 de la Unidad Educativa. En la actualidad, se cuenta con 69 estudiantes 40 con discapacidad auditiva y 29 con discapacidad visual, desde estimulación temprana hasta décimo año de educación básica. En base a una entrevista con la Directora y Subdirectora de la Unidad Educativa, se toma como población a los 40 estudiantes con discapacidad auditiva, determinando que sería de gran apoyo utilizar alternativas tecnológicas en el proceso de enseñanza-aprendizaje de la lengua de señas en los niños de 7 a 15 años. La área de estudio con la cual se va a trabajar es la asignatura de Lengua y Literatura dentro de la lectoescritura, teniendo en cuenta, que las señas, gestos e imágenes, es el principal modo de comunicación entre las personas con discapacidad auditiva, razón por la cual, esta investigación establece como objeto de estudio a los 10 estudiantes de: Segundo de Básica (1 estudiante), Cuarto de Básica (5 estudiantes) y Quinto de Básica (4 estudiantes) para el desarrollo del software que con uso de dispositivos de reconocimiento de voz traduzca a lenguaje de señas.

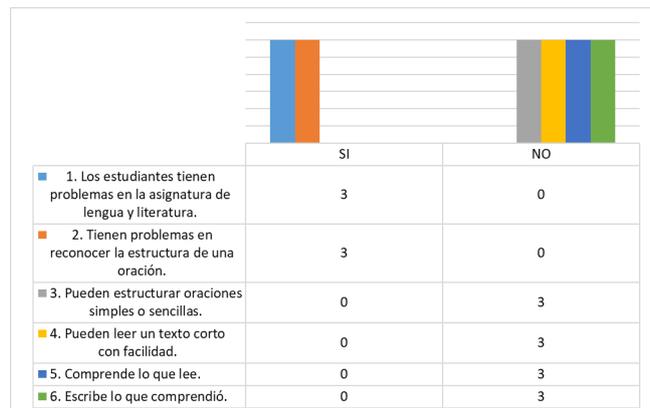
Luego de la determinación del objeto de estudio se realizó un diagnóstico del porcentaje de discapacidad de los 10 estudiantes según su diagnóstico clínico del carnet del CONADIS como se puede ver en la Fig. 8.

Fig. 8. Porcentaje de discapacidad auditiva de cada estudiante.



Asimismo, se realizaron 7 interrogantes para determinar las necesidades instruccionales que tienen los niños al momento de la enseñanza-aprendizaje del lenguaje de señas como se puede ver en la Fig. 9:

Fig. 9. Resultados de las interrogantes realizadas a los docentes.



Con referencia a las Fig. 9 se evidencia que todas las preguntas realizadas a los docentes del Segundo, Cuarto y Quinto de Básica obtienen una puntuación de 3, que es el 100 % en cada pregunta. Además, los docentes supieron manifestar que:

- La escritura y lectura lo hacen con ayuda y luego de que ha sido explicado varias veces en lengua de señas.
- Falta de conocimiento del lenguaje de señas.
- No hay lectura comprensiva.
- No escriben correctamente con las reglas gramaticales.
- No hay memoria de largo plazo.
- En la lengua de señas la estructura es diferente de la gramática española.

1. Requerimientos Administrador

- 1) **R1.** Realizar el mantenimiento de los registros del Sistema, a través del ingreso de la información general de

los estudiantes, categorías, subcategorías, vocabulario, diccionario, oraciones simples y juegos, ver Fig. 10.

2. *Requerimientos Profesor*

- 1) **R2.** Registrar la planificación de las actividades por cada estudiante para un periodo de tiempo establecido, ver Fig. 11.
- 2) **R3.** Evaluar el aprendizaje alcanzado por cada estudiante a través del reporte de estudiantes para su retroalimentación, ver Fig. 11.

3. *Requerimientos Estudiante*

- 1) **R5.** Realizar la ejecución de las actividades planificadas por parte de los docentes, ver Fig. 12.

4. *Diagramas de Casos de Uso*

Un Diagrama de Caso de Uso muestra la relación entre los actores y los casos de uso del sistema. Representa la funcionalidad que ofrece el sistema en lo que se refiere a su interacción externa [28].

Fig. 10. Diagrama de Caso de Uso del Administrador del Sistema.

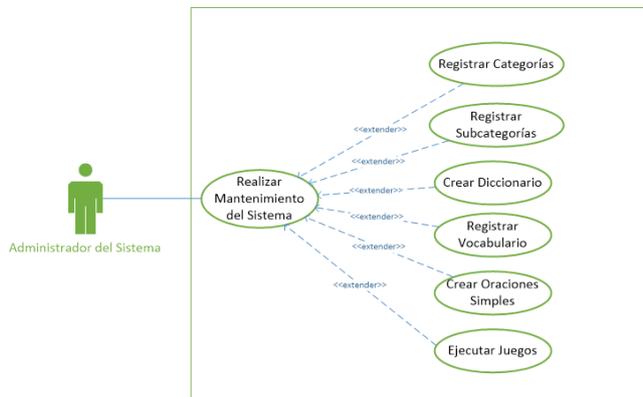


Fig. 11. Diagrama de Caso de Uso de los Profesores.

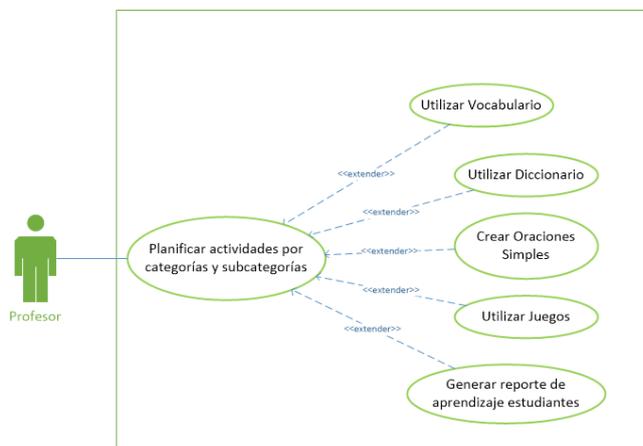
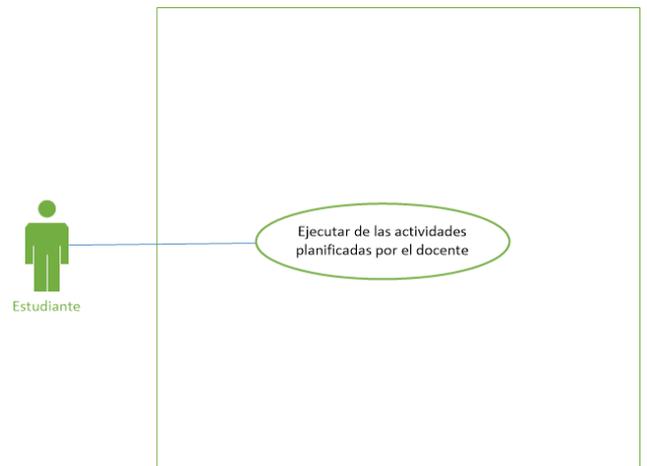


Fig. 12. Diagrama de Caso de Uso de los Estudiantes.



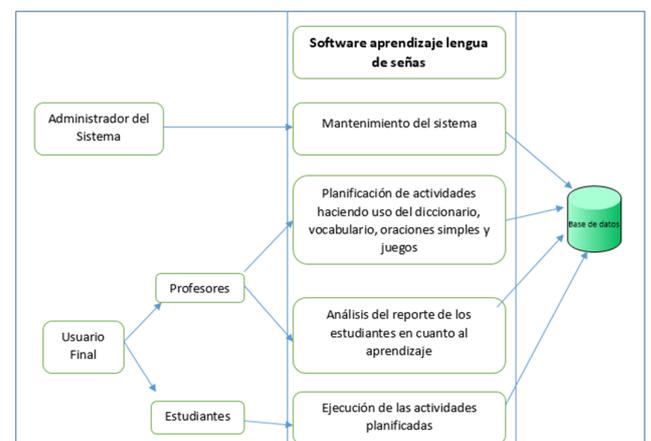
Los diagramas de casos de uso permiten visualizar la relación de cada actor con los procesos del sistema.

B. *Diseño de software*

1. *Arquitectura*

La arquitectura del software para el aprendizaje de la lengua de señas se puede ver en la Fig. 13, el mismo que esta diseñado en tres capas:

Fig. 13. Diagrama de Bloques de la Arquitectura del Software de aprendizaje de la lengua de señas.

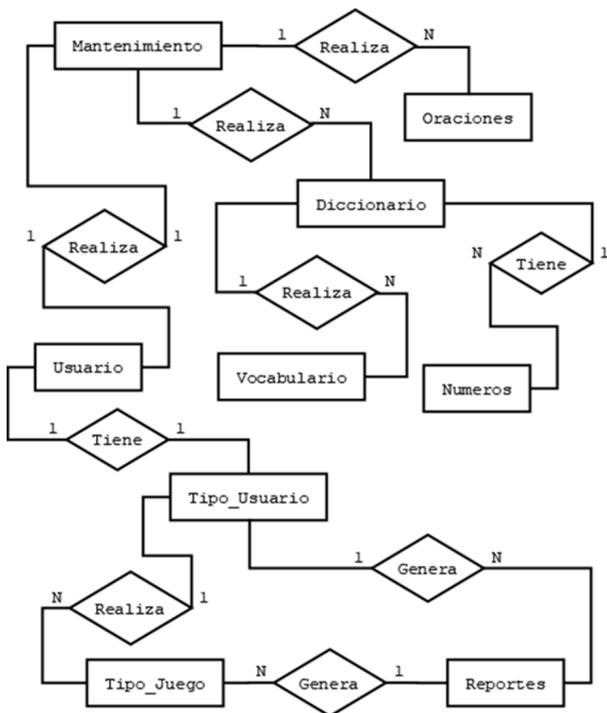


- **Capa de Presentación:** Se encarga de que el sistema interactúe con el usuario y viceversa: interfaces de administración del sistema, profesor y estudiante.
- **Capa de Negocio:** Permite la funcionalidad del sistema a través del mantenimiento del sistema, planificación de actividades haciendo uso del diccionario, vocabulario, oraciones simples y juegos, análisis de los reportes de los estudiantes en cuanto al aprendizaje y ejecución de las actividades programadas.
- **Capa de Datos:** Esta capa permite el almacenamiento de los datos del sistema y de los usuarios en una base de datos local.

2. Modelo entidad relación

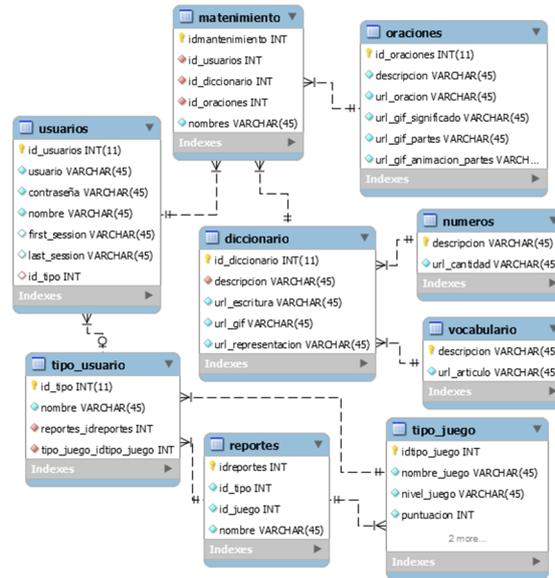
El modelo de datos de entidad-relación (ER) se basa en una percepción de un mundo real que consiste en un conjunto de objetos llamados entidades y las relaciones entre estos objetos. A continuación, en la Fig. 14 se puede observar el modelo entidad relación del Software de aprendizaje de la lengua de señas:

Fig. 14. Modelo entidad relación del Software de aprendizaje de la lengua de señas.



Una base de datos es un conjunto de tablas o entidades que tiene una estrecha relación entre sí y contienen atributos con sus respectivos tipos de datos, en donde se almacenarán los datos manipulados en el sistema, como se puede ver en la Fig. 15:

Fig. 15. Base de Datos del Software de aprendizaje de la lengua de señas.



La Base de Datos del software de aprendizaje de la lengua de señas será alimentada por los usuarios administrador, docentes y estudiantes.

C. Implementación y pruebas unitarias

La construcción del software fue realizada por estudiantes de la carrera de Ingeniería de Sistemas de la Universidad Católica de Cuenca. Tomando en consideración, los resultados obtenidos del estudio preliminar y la parte instruccional de los estudiantes de la Unidad Educativa Especial Claudio Neira Garzón, se desarrolló el software utilizando el lenguaje de programación Java (NetBeans), MySQL (Base de Datos) y Flash e Illustrator (diseño gráfico).

El sistema cuenta con 2 tipos de usuarios:

- Administrador: es quien se encarga de la alimentación de información en las categorías contenidas en el programa (guardar, actualizar y eliminar).
- Usuario final: se divide en 2 tipos: profesores y estudiantes, son aquellos que harán uso de las distintas funciones del sistema.

El sistema permite al usuario la enseñanza-aprendizaje de la lectoescritura, tanto en el lenguaje español como en el lenguaje de señas, el mismo que está conformado por varios formularios, entre sus formularios principales están: diccionario, vocabulario y oraciones simples. Esta aplicación tiene una base de datos en la cual se almacena las imágenes, gifs e información de los diferentes formularios.

Las categorías y subcategorías fueron elaboradas en base al DICCÍ SEÑAS de la Pontificia Universidad Católica de Chile publicado en sus Tecnologías de Inclusión CEDETI <http://diccisenas.cedeti.cl/AR/categoria/11/> y el Lenguaje de Señas Guía Básica sobre una Comunicación Especial – Tomo I de la página Sordos Ecuador <http://www.sordosecuador.com/p/lenguaje-de-senas-guia-basica.html>.

La página principal del sistema se puede ver en la Fig. 16:

Fig. 16. Pantalla inicial del sistema.



1. Creación de oraciones simples

Para la creación de una oración simple se utiliza el formulario oraciones simples, el mismo que existe tanto para docentes como estudiantes y se apoya en los formularios de diccionario y vocabulario en caso de no saber el significado de alguna palabra, así mismo, se tiene juegos que permiten reforzar el aprendizaje del diccionario, vocabulario y oraciones simples; también hay un formulario de reportes de estudiantes que permite dar un seguimiento del aprendizaje de los mismos.

La estructura gramatical del lenguaje de señas tiene una interpretación diferente al momento de ser traducida al lenguaje español, su estructura gramatical es la siguiente:

SUJETO + PREDICADO + VERBO

(Si la oración tiene tiempo, es decir está en pasado o en futuro sería de la siguiente forma)

TIEMPO + SUJETO + PREDICADO + VERBO

El sistema busca corregir las deficiencias gramaticales y mejorar el léxico del lenguaje español y optimizar el aprendizaje en los estudiantes de la lectoescritura a través del diccionario, vocabulario, juegos y oraciones simples.

Fig. 17. Juego que permite crear una oración simple.



2. Página del profesor

El profesor a través de un micrófono puede introducir una palabra que es capturada mediante el Talking y com-

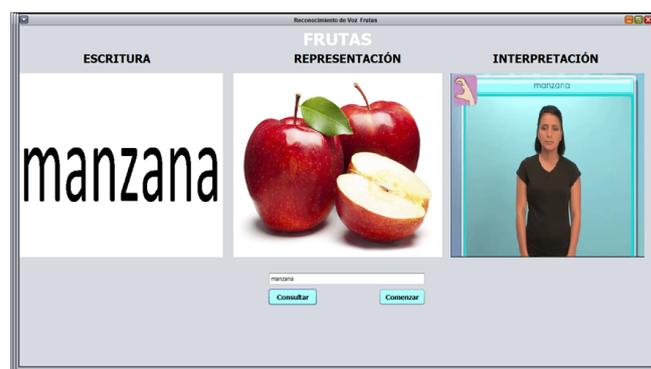
para con su base datos, esta palabra capturada es también comparada con una palabra almacenada en una base de datos MySQL que permite la visualización de la escritura, representación y el significado de la palabra en lenguaje de señas.

El profesor puede hacer uso del diccionario, vocabulario, oraciones simples y reportes de estudiantes:

- El diccionario tiene información de las subcategorías: alimentos y bebidas, acciones y cualidades, escuela, hogar, días y meses, ocupaciones y términos lingüísticos.
- El vocabulario tiene información de las categorías: alimentos y bebidas, escuela, hogar, días y meses y ocupaciones.
- Las oraciones simples tienen información de: alimentos y bebidas, acciones y cualidades, escuela, hogar y términos lingüísticos.
- Reporte de estudiantes: ayuda al profesor a obtener una visión del aprendizaje y rendimiento de sus estudiantes con respecto al aprendizaje de los contenidos de la asignatura a través del uso del programa, además, permite generar estadísticas de los tiempos que los estudiantes emplean para realizar cada actividad y retroalimentar su instrucción, gracias a las mismas se puede saber también, los contenidos que los estudiantes están con mayores dificultades en su aprendizaje.

Dentro de esta página, el profesor enseña la lectoescritura del lenguaje español y el lenguaje de señas de una forma más entretenida y dinámica. Ya que en la misma, se muestra la Escritura (Lenguaje Español), la Representación (Imagen) y la Interpretación (Lengua de Señas), como se puede ver en la Fig. 18.

Fig. 18. Pantalla de presentación del vocabulario.



Además, los profesores en el caso que tengan una discapacidad auditiva podrán hacer uso del teclado para la utilización del sistema con todas sus funcionalidades.

3. Página del estudiante

El estudiante a través de esta página puede hacer uso de las siguientes opciones: diccionario, vocabulario y oraciones simples:

- El diccionario tiene un juego de memoria que permite encontrar las palabras o imágenes pares.
- El vocabulario tiene un juego que permite unir un artículo con un sustantivo.
- Las oraciones simples tienen un juego que permite completar una oración.

Todos los juegos tienen 3 niveles de complejidad, existe un nivel básico, intermedio y avanzado; la información presentada en los juegos, son recuperados desde una base de datos y mostrados de forma aleatoria en cada uno de los formularios. Los mismos, que permitirán reforzar el aprendizaje en los estudiantes.

D. Integración y pruebas del sistema

El sistema ha sido integrado de acuerdo a los requerimientos de los usuarios y probado de forma continua a través de secciones programadas con los estudiantes y los docentes, lo que ha permitido realizar cambios significativos y obtener un sistema muy amigable para con los estudiantes y docentes. Algunas de sus apreciaciones fueron:

- La interfaz gráfica de la aplicación es del agrado de los estudiantes, docentes y muy intuitiva en su manipulación.
- El uso de los juegos no presenta dificultad en los estudiantes, lo pudieron realizar con indicaciones cortas.
- Su uso en el aula permite complementar los aprendizajes de los estudiantes y las explicaciones del profesor.
- La herramienta fomenta el desarrollo de habilidades de pensamiento (crear, analizar y evaluar).
- Al realizar las pruebas a un niño con discapacidad auditiva se notó que mejoró su habilidad de memoria y gramática del lenguaje español reduciendo sus tiempos en los juegos y ordenando de mejor manera las oraciones.

Sin embargo, el sistema está en su primera etapa de desarrollo, al mismo que se seguirá alimentando más funcionalidades que le permita abordar los contenidos en su totalidad de la asignatura de Lengua y Literatura.

E. Operación y mantenimiento

Se ha registrado información del diccionario, vocabulario, oraciones simples y juegos:

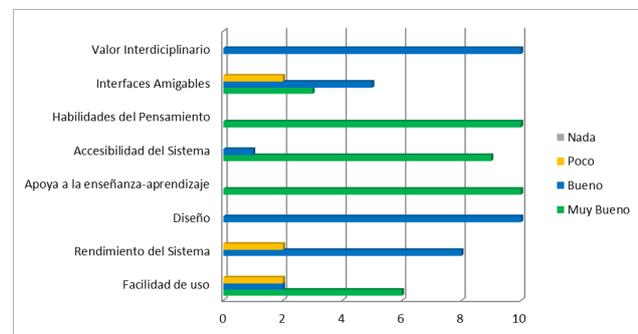
- En el diccionario se registró información de las subcategorías: alimentos y bebidas, acciones y cualidades, escuela, hogar, días y meses, ocupaciones y términos lingüísticos.
- En el vocabulario se registró información de las categorías: alimentos y bebidas, escuela, hogar, días y meses y ocupaciones.
- En las oraciones simples se registró información de: alimentos y bebidas, acciones y cualidades, escuela, hogar y términos lingüísticos. Además, las oraciones simples se complementa con un juego que permite completar una oración.

La operación y mantenimiento del Software de aprendizaje de la lengua de señas es un proceso continuo que se realiza de acuerdo a los contenidos de la asignatura de Lengua y Literatura, material didáctico y las necesidades educativas de los profesores y estudiantes. Además se pueden crear nuevas categorías y subcategorías.

V. DISCUSIÓN

La evaluación del sistema con los estudiantes, a través de una ficha de observación, de la facilidad de uso del software y su utilidad en el proceso de enseñanza-aprendizaje de la lectoescritura, arroja los resultados que se puede ver en la Fig. 19:

Fig. 19. Resultados de la ficha de observación aplicada a los estudiantes en el momento de utilizar el sistema.



Esto determina que el 47 % del sistema es Muy Bueno e interactivo de acuerdo a la ficha de observación aplicada a los estudiantes en el momento de utilizar el sistema, asimismo el 45 % es Bueno, el 8 % Poco y el 0 % Nada. Haciendo un análisis se interpreta que el sistema contribuye con las habilidades del pensamiento (crear, analizar y evaluar), apoya el proceso de enseñanza-aprendizaje, es accesible para niños con discapacidad auditiva (traduce de voz a lengua de señas), es fácil de usar y sus interfaces son amigables.

VI. CONCLUSIONES

Se concluye que el sistema fue desarrollado en su primera etapa, el mismo que apoya el proceso de enseñanza-aprendizaje de los niños con discapacidad auditiva en la asignatura de Lengua y Literatura de la Unidad Educativa Claudio Neira Garzón dentro de la lectoescritura, los formularios diseñados en el sistema son: Profesor, Estudiante y Administrador. Los formularios del Profesor y Estudiante tienen formularios de diccionario, vocabulario, oraciones simples y juegos. Los juegos tienen tres niveles de complejidad y el formulario del Profesor además tiene el reporte de estudiantes.

En las pruebas realizadas con el sistema se han obtenido muy buenos resultados, teniendo una excelente acogida por parte de los docentes y estudiantes que demostraron mucho interés por usar el software para la enseñanza-aprendizaje, en el análisis del uso del programa de los niños arroja

buenos resultados y también sugerencias de cambios en el mismo para mejorar su interactividad con los estudiantes y profesores. Además, a través del uso del sistema los niños incrementaron el nivel de comprensión y entendimiento del lenguaje español y la lengua de señas.

Finalmente, el software que con el uso de un dispositivo de reconocimiento de voz traduce a lenguaje de señas sirve como herramienta de trabajo en la labor educativa de la Unidad Educativa Especial Claudio Neira Garzón en la asignatura Lengua y Literatura, permitiendo la creación de oraciones simples de forma interactiva y dinámica. Este software está orientado principalmente a niños entre 7 y 15 años con déficit auditivo parcial o total, además, beneficia a personas adultas y familiares de las personas con déficit auditivo parcial o total, entre otras personas.

REFERENCIAS

- [1] T. Constitucional, "Constitución de la república del Ecuador," *Quito-Ecuador: Registro Oficial*, vol. 449, pp. 20–10, 2008.
- [2] S. N. de Planificación y Desarrollo, "Plan nacional de desarrollo 2017-2021. toda una vida.," *Ecuador*, Septiembre 2017.
- [3] CONADIS, "Normas jurídicas en discapacidad Ecuador.," Mayo 2014.
- [4] B. MUNDIAL, "Informe mundial la discapacidad," *Educación*, vol. 218, no. 219, p. 219, 2011.
- [5] CONADIS, "Información estadística de personas con discapacidad.," Marzo 2018.
- [6] M. A. C. Molina, J. K. V. Durán, and D. A. R. Espinosa, "Accesibilidad web en las instituciones de salud de la ciudad de Cuenca. análisis preliminar," in *IV Conferencia Internacional sobre Aplicación de Tecnologías de la Información y Comunicaciones para mejorar la Accesibilidad (ATICAcces 2016)*, pp. 125–132, Servicio de Publicaciones, 2016.
- [7] J. K. V. Durán, M. A. C. Molina, and D. X. P. Japón, "Funciones de accesibilidad que logran y mejoran las competencias de uso y manejo de los teléfonos inteligentes," in *CIIEE 2017*, Marzo 2017.
- [8] M. A. C. Molina, J. K. V. Durán, and V. S. E. Jara, "Accesibilidad web en las instituciones de salud de la ciudad de Cuenca. análisis preliminar," in *VIII Congreso Internacional sobre Aplicación de Tecnologías de la Información y Comunicaciones Avanzadas (ATICA2017)*, *V Conferencia Internacional sobre Aplicación de Tecnologías de la Información y Comunicaciones para mejorar la Accesibilidad*, *I Congreso Internacional Educar en la Sociedad Red: Realidad, Retos y Perspectivas*, pp. 141–148, Servicio de Publicaciones, 2017.
- [9] CONADIS, "Causas de discapacidad," Mayo 2013.
- [10] M. Figueroa Cruz, G. Vázquez Zubizarreta, and M. A. Campoverde Molina, "Software educativo para el desarrollo de habilidades de la conducta adaptativa en personas con discapacidad intelectual," *VARONA ISSN: 1992-8238*, no. 61, p. 11, 2015.
- [11] R. Claros-Kartchner, "La inclusión de las personas sordas, como grupo étnico, en los sistemas educativos," 2017.
- [12] A. Ecuador, "Ley orgánica de discapacidades," *Registro Oficial N*, vol. 796, Septiembre 2012.
- [13] M. P. Godoy, S. González, and M. Verdugo, "Guía de apoyo técnico-pedagógico: Necesidades educativas especiales en el nivel de educación parvularia," *Santiago de Chile: Editorial Atenas*, p. 46, 2008.
- [14] S. Cano, J. Muí, C. Collazos, V. B. Amador, et al., "Aplicación móvil para el aprendizaje de la lectoescritura con fitzgerald para niños con discapacidad auditiva," in *Anais dos Workshops do Congresso Brasileiro de Informática na Educação*, vol. 4, p. 240, 2015.
- [15] X. C. Alejandro Oviedo and R. Cabezas, "Ecuador, atlas sordo," 2015.
- [16] C. Becerra Sepúlveda, "Lenguaje y educación en niños sordos: encuentros y desencuentros," *REXE. Revista de Estudios y Experiencias en Educación*, no. 14, 2008.
- [17] D. L. Parra and M. Luque-Rojas, "Necesidades específicas de apoyo educativo del alumnado con discapacidades sensorial y motora," *Summa psicológica UST (En línea)*, vol. 10, no. 2, pp. 57–72, 2013.
- [18] E. R. Linares, "El aprendizaje de la lectoescritura en los niños y niñas sordos," *Caleidoscopio, Revista digital de contenidos educativos*, no. 2, p. 1, 2009.
- [19] S. E. A., "El podcasts en la enseñanza de la lectoescritura," *Revista Virtual Universidad Católica del Norte*, pp. 9–26, 2012.
- [20] R. Montealegre and L. A. Forero, "Desarrollo de la lectoescritura: adquisición y dominio," *Acta colombiana de psicología*, vol. 9, no. 1, pp. 25–40, 2006.
- [21] M. Gértrudix, S. Álvarez, A. Galisteo, M. del Carmen Gálvez, and F. Gértrudix, "Acciones de diseño y desarrollo de objetos educativos digitales: programas institucionales," *RUSC. Universities and Knowledge Society Journal*, vol. 4, no. 1, 2007.
- [22] M. de Educación, "Circular nro. mineduc-ve-2016-00001-c," p. 12, Agosto 2016.
- [23] S. B. Stainback, "L'educació inclusiva: definició, context i motius," *Suports: revista catalana d'educació especial i atenció a la diversitat*, vol. 5, no. 1, pp. 18–25, 2001.
- [24] "Educación inclusiva," 2006.
- [25] G. Echeita Sarrionandia and C. Duk Homad, "Inclusión educativa," *REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, vol. 6, no. 2, 2008.
- [26] M. Granada Azcárraga, M. P. Pomés Correa, and S. Sanhueza Henríquez, "Actitud de los profesores hacia la inclusión educativa," *Papeles de trabajo-Centro de Estudios Interdisciplinarios en Etnolingüística y Antropología Socio-Cultural*, no. 25, pp. 0–0, 2013.

- [27] V. Esteller and E. Medina, "Procesos de desarrollo de software y materiales educativos computarizados," *Eduweb, Revista de Tecnología de Información y Comunicación en Educación*, vol. 6, no. 1, 2012.
- [28] X. F. Grau and M. I. S. Segura, "Desarrollo orientado a objetos con uml," *Recuperado el*, vol. 1, 2008.

Recibido: 3 de mayo de 2018

Aceptado: 15 de junio de 2018



Evaluación de seguridad de la información en las páginas web pertenecientes a los municipios de la provincia del Cañar

Cristhian Flores Urgilés^{1*}, Beatriz Zhinin Aguayza¹, Alexandra Segovia Cantos¹, Mercedes Mayancela Zhinin¹,
Jessica Marlene García¹

Ingeniería en Sistemas, Universidad Católica de Cuenca extensión Cañar¹

*chfloresu@ucacue.edu.ec

DOI: https://doi.org/10.26871/killkana_tecnica.v2i1.286

Resumen

La presente investigación se desarrolló con el fin de buscar el aseguramiento de la información de las páginas web pertenecientes a las municipalidades de la provincia del Cañar mediante un análisis de vulnerabilidades, buscando prevenir las amenazas que se puedan presentar y estas a la vez pongan en peligro la estabilidad de dichas páginas. Se realizó una investigación teórica sobre aseguramiento de aplicativos webs, estándares y buenas prácticas de seguridad de la información y las diferentes herramientas de escáner de vulnerabilidades de páginas web, para finalmente luego del análisis respectivo utilizar la herramienta Acunetix 11, además tomando como referencia el Top Ten de Owasp 2017. Como resultado se obtuvo un informe detallado de las vulnerabilidades encontradas durante el proceso de escaneo de las diferentes páginas municipales; para finalizar con las respectivas recomendaciones de seguridad para las amenazas más comunes.

Palabras clave: Owasp, Vulnerabilidad, informática, seguridad de información, amenaza informática, escaneo de vulnerabilidades.

Abstract

The present research was developed in order to seek the security of the information in the web pages of the municipality from the Cañar Province through a vulnerability analysis, seeking to prevent the threats that they can present and these can risk the stability of such web pages. First, a theoretical investigation about the security of web applications, standards, good information security practices, and the different vulnerability scan tools of web pages was carried out. Finally, after the respective analyses the Acunetix 11 tool was used, taking as reference the OWASP 2017 Top Ten security risks. As a result, a detailed report about the vulnerabilities that were found during the scanning process of the different municipal web pages was obtained, to conclude with the respective security recommendations for the most common threats.

Key words: OWASP, Informatics Vulnerability, information security, computer threats, vulnerability scanning.

I. INTRODUCCIÓN

En la actualidad la mayoría de las empresas públicas y privadas utilizan sitios web y aplicativos móviles como herramientas fundamentales para informar sus actividades o para automatizar algunos de sus procesos administrativos y operativos; en vista de que los incidentes de seguridad en este caso involucrarían la confidencialidad, integridad o disponibilidad de todos los datos asociados al aplicativo web, así como también a los usuarios del mismo.

Consientes que desde el primer momento en que un servidor, computadora, se conecta a internet, se encuentra expuesto a diferentes tipos de amenazas informáticas, la mayoría derivada de vulnerabilidades propias tan simples como la versión de sistemas operativos u otros más complejos, en vista que los cambios tecnológicos ocurren rápidamente, lo que implica que a las actualizaciones de los diferentes programas, sistemas operativos y controladores

de hardware se tengan que realizarse periódicamente, caso contrario se convertirán en vulnerabilidades que podrían ser explotadas.

En la actualidad se acostumbra a utilizar referentes o buenas prácticas de seguridad de aplicativos webs como una forma de analizar los riesgos más comunes a los que los aplicativos se encuentran expuestos.

Mediante software que escanean riesgos de seguridad de aplicativos webs se permiten obtener reportes, comparándolos con los estándares, que servirán para determinar los riesgos existentes, para así implementar salvaguardas que permitirán minimizarlo.

II. MATERIALES Y MÉTODOS

A. Selección de la herramienta

La herramienta a utilizar para el escaneo de vulnerabilidades de los aplicativos webs fue Acunetix 11. Debido a

que dicha herramienta es capaz de realizar un escaneo a cualquier sitio web a través del protocolo HTTP/HTTPS. Cabe resaltar que esta herramienta está actualizada a la última versión del Top ten de Owasp que es la del 2017.

Otra de las razones del porque se seleccionó la herramienta Acunetix 11 es que, al momento de terminar el análisis de vulnerabilidades de las páginas web, emite un informe detallado de cada uno de los riesgos informáticos tomando como referencia al Top Ten de Owasp 2017, encontrados en los aplicativos webs analizados.

B. Población y muestra

Para el análisis de las páginas se realizó una investigación del número de cantones que tiene la provincia de Cañar.

La provincia de Cañar está conformada por 7 cantones de las cuales 6 de ellas tienen página Web solamente el cantón Suscal no posee página web.

Luego de haber realizado el análisis a cada una de ellas se llegó a obtener el resultado que a continuación se detallara.

C. Metodología

- Como primer paso fue la identificación del dominio URL, de cada uno de los municipios que poseen sitios web.
- Luego de un tiempo de espera que es aproximadamente un día, tiempo que depende del contenido de cada una de las páginas.
- Una vez terminado el análisis emite un informe en donde se presenta todas las vulnerabilidades que tiene la página de acuerdo al Top Ten de Owasp 2017.
- En el siguiente paso Se procederá a la tabulación de los datos obtenidos para encontrar los riesgos informáticos más comunes.
- En base a los resultados del punto anterior se presentará una propuesta de prevención para los riesgos informáticos más comunes.

III. DESARROLLO

A. Seguridad de la información

La seguridad informática consiste en proteger adecuadamente la información almacenada en un medio digital o no, en base a técnicas, estándares, medidas preventivas que nos ayuden a garantizar la protección y aseguramiento de esa información, el cual resulta ser el bien más valioso para las organizaciones que lo manejan. Permitiendo de esta manera garantizar la integridad, confidencialidad y disponibilidad de dicha información conocidos como los pilares fundamentales para asegurar la información, dichos conceptos en una aplicación web serían aplicados de la siguiente forma:

- 1) Confidencialidad es la propiedad que asegura que solo los que están autorizados tendrán acceso a la información.

- 2) La integridad es la propiedad que asegura la no alteración de la información.
- 3) La autenticación es la propiedad que hace referencia a la identificación. Es el nexo de unión entre la información y su emisor.
- 4) Disponibilidad La información estará disponible en cualquier momento solo a usuarios autorizados, evitando la divulgación.

Sin embargo, hoy en día no se puede ofrecer una protección absoluta de la información ya que cada vez existen nuevas amenazas, riesgos a las que se vuelven vulnerables los sistemas de información.

B. Seguridad en los aplicativos web

Las aplicaciones web vienen acompañadas de una variedad de vulnerabilidades de seguridad. Es por ello que las organizaciones se debe aplicar estándares, normas y reglas que ayuden a controlar a que la información y recursos estén protegidos.

1. Vulnerabilidades

Probabilidades que existen de que una amenaza se materialice contra un activo. No todos los activos son vulnerables a las mismas amenazas. Por ejemplo, los datos son vulnerables a la acción de los hackers, mientras que una instalación eléctrica es vulnerable a un cortocircuito. Al hacer el análisis de riesgos hay que tener en cuenta la vulnerabilidad de cada activo.[1]

2. Amenazas

Son agentes que poseen la cualidad de aprovechar una debilidad o vulnerabilidad para llevar a cabo un acto que afecte a un sistema. Las amenazas son directamente proporcionales a las falencias que tenga un sistema, si este posee grandes debilidades en su modo de operar o diseño, las amenazas serán altas y tendrán un gran impacto en el mismo.[1]

Amenazas más comunes para aplicativos webs

En la actualidad los aplicativos Web se encuentran en constante amenaza debido a que los ataques son cada vez más sofisticados. El Internet es el medio principal de ataque ya que, si no se ejecuta las medidas necesarias de seguridad, la información se vuelve vulnerable.

Entre los más comunes se puede citar los siguientes:

3. Inyección

Las fallas de inyección, como SQL, NoSQL, OS e inyección LDAP, ocurren cuando los datos que no son de confianza se envían a un intérprete como parte de un comando o consulta. Los datos hostiles del atacante pueden engañar al intérprete para que ejecute comandos no deseados o acceda a datos sin la debida autorización.

4. *Cross site scripting*

Cross-site Scripting (XSS) se refiere al ataque de inyección de código del lado del cliente en el que un atacante puede ejecutar scripts maliciosos en un entorno legítimo. Sitio web o aplicación web. XSS ocurre cuando una aplicación web utiliza una entrada de usuario no validada o no codificada dentro de la salida genera

5. *Broken Access control*

Las restricciones sobre lo que los usuarios autenticados pueden hacer a menudo no se aplican correctamente. Los atacantes pueden explotar estos defectos para acceder a funcionalidades y / o datos no autorizados, como acceder a cuentas de otros usuarios, ver archivos confidenciales, modificar datos de otros usuarios, cambiar derechos de acceso, etc.

6. *Broken authentication*

Las funciones de aplicación relacionadas con la autenticación y la administración de sesión a menudo se implementan incorrectamente, lo que permite a los atacantes comprometer contraseñas, claves o tokens de sesión, o explotar otras fallas de implementación para asumir.

7. *Sensitive data exposure*

Muchas aplicaciones web y API no protegen adecuadamente los datos confidenciales, como financieros, de salud. Los atacantes pueden robar o modificar dichos datos protegidos débilmente para realizar fraude con tarjetas de crédito, robo de identidad u otros delitos. Los datos confidenciales pueden verse comprometidos sin protección adicional, como el cifrado en reposo o en tránsito, y requieren precauciones especiales cuando se intercambian con el navegador.

C. *Análisis de Riesgos*

Se denomina riesgo a la posibilidad de que se materialice o una amenaza aprovechando una vulnerabilidad. No constituye riesgo una amenaza cuando no hay vulnerabilidad ni una vulnerabilidad cuando no existe amenaza para la misma.[2]

El análisis de riesgo consiste en la cuantificación de dicho riesgo en base a la probabilidad de ocurrencia conjuntamente con el impacto que causaría al sistema la materialización de la amenaza.

D. *Estándares y Buenas prácticas de Seguridad en aplicativos Web.*

1. *Owasp*

Es un proyecto de código abierto destinada a la de seguridad en aplicaciones Web, esta comunidad está dedicada a habilitar a las organizaciones para desarrollar, comprar y mantener aplicaciones confiables.[3]

¿Qué es el Owasp top 10?

El objetivo del proyecto Top 10 es crear conciencia sobre la seguridad de las aplicaciones mediante la identificación de algunos de los riesgos más críticos que enfrentan las organizaciones. Algunos estándares, libros, herramientas y organizaciones hacen referencia al proyecto Top 10, que incluye MITRE, PCI DSS, DISA, FTC y muchos más.[4]

2. *Wasc (Web Application Security Consortium)*

Los miembros Web Application Security Consortium, han unido esfuerzos de cooperación para desarrollar y promover un estándar de organización y clasificación de amenazas de seguridad web. Lo cual provee de un lenguaje común a desarrolladores, fabricantes de software, personal de seguridad y auditores, para tratar aspectos afines con la seguridad.[3]

Facilita un entendimiento profundo sobre los riesgos que amenazan a las aplicaciones Web. Wasc publica constantemente una guía información útil.

3. *Herramientas de escáner de vulnerabilidades en aplicativos webs.*

Estas herramientas permiten hacer un análisis a las páginas web para detectar las vulnerabilidades y riesgos de seguridad a las que están expuestas. Los escáneres de vulnerabilidad pueden ser de pago o gratuitos.

4. *Grabber*

Es unas herramientas más de escaneo de vulnerabilidades de las aplicaciones web, realizando exploraciones en las que se puede detectar amenazas como: SQL injection, Ajax Testing, etc. Se utiliza para analizar sitios web de cualquier tamaño debido a su fácil manejo.

5. *Vega*

Se puede utilizar para encontrar inyección SQL, inyección de cabecera, listado de directorios, inyección cáscara, Cross site scripting, la inclusión de archivos y otras vulnerabilidades de las aplicaciones web. Esta herramienta también se puede ampliar mediante una potente API desarrollada en JavaScript.[5]

6. *Wapiti*

Realiza escaneos de "caja negra"(no estudia el código fuente) de la aplicación web al rastrear las páginas web de la aplicación web desplegada, en busca de scripts y formularios donde puede inyectar datos.[6]

7. *W3af*

W3af es un Marco de auditoría y ataque de aplicaciones web. El objetivo del proyecto es crear un marco para ayudarlo a proteger sus aplicaciones web al encontrar y explotar todas las vulnerabilidades de las aplicaciones web.[7]

8. WebScarab

WebScarab es un marco para analizar aplicaciones que se comunican mediante los protocolos HTTP y HTTPS. Está escrito en Java y, por lo tanto, es portátil para muchas plataformas. WebScarab tiene varios modos de operación, implementados por una cantidad de complementos. En su uso más común, WebScarab funciona como un proxy de interceptación, lo que permite al operador revisar y modificar las solicitudes creadas por el navegador antes de enviarlas al servidor, y revisar y modificar las respuestas devueltas desde el servidor antes de que el navegador las reciba.[8]

9. Sqlmap

Esta herramienta permite encontrar automáticamente las vulnerabilidades de sql Injection en la base de datos, con la que se encuentre trabajando el sitio web.

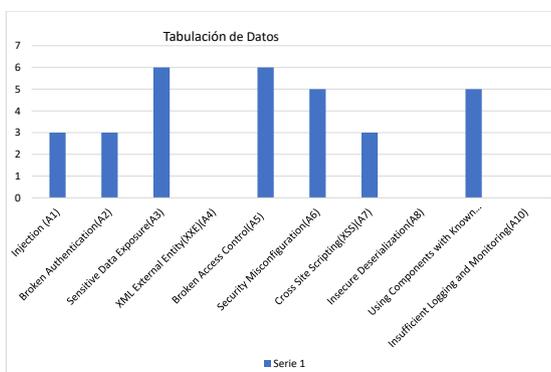
IV. DISCUSIÓN

Luego de haber culminado el proceso de escaneo de vulnerabilidades utilizando la herramienta de Acunetix 11, en base al estándar Owasp Top Ten 2017 se obtuvo los siguientes resultados.

TABLA I. Vulnerabilidades encontradas

VULNERABILIDADES	TOTAL
Injection (A1)	3
Broken Authentication(A2)	3
Sensitive Data Exposure(A3)	6
XML External Entity(XXE)(A4)	0
Broken Access Control(A5)	6
Security Misconfiguration(A6)	5
Cross Site Scripting(XSS)(A7)	3
Insecure Deserialization(A8)	0
Using Components with Known Vulnerabilities(A9)	5
Insufficient Logging and Monitoring(A10)	0

Fig. 1. Ilustración del rango de vulnerabilidades



Fuente: Autor del proyecto

De acuerdo con los resultados que se muestran en la fig.1 se logró determinar las vulnerabilidades que son más perjudiciales en una página web, entre ellas están:

A. Broken access control

Esta vulnerabilidad es la que comúnmente se encuentra en las páginas web. Las restricciones sobre lo que los usuarios autenticados pueden hacer a menudo no se aplican correctamente, es decir que un usuario puede hacer clic en un enlace no deseado y dar paso a un atacante, para que este pueda acceder ya sea a cuentas de otros usuarios, ver y modificar Datos de usuarios, tomar control de la computadora. Por lo que el sitio web podría estar en riesgo de un ataque de clickjacking.

1. Protección

Para que una página este protegido ante esta vulnerabilidad se recomienda a los desarrolladores, realizar configuraciones tales como:

Frameguard: configura cabeceras X-Frame-Options en modo DENY ayuda a que no sea atacado por Clickjacking.

HidePoweredBy: desactiva cabeceras X-Powered-By evita que el atacante obtenga información.

XSS Filter: activa la cabecera X-XSS-Protection en modo "1; mode=block" para que el navegador puede detectar ataques XSS y los mismos puedan ser evitados.[9]

B. Sensitive data exposure

Esta vulnerabilidad es también una de las más comunes en páginas, permite a un atacante acceder a los datos confidenciales que no se encuentran protegidas, de esta manera pueden modificar dichos datos para realizar fraudes, robo de identidad.[3]

1. Protección

Una de las protecciones que se debe realizar para prevenir esta vulnerabilidad es clasificando datos procesados almacenados y transmitidos por un sistema. Aplicar controles según la clasificación.

No almacenar datos confidenciales innecesariamente hay que desecharlo una vez ya utilizado.

Existen vulnerabilidades que pocas páginas poseen, las cuales son:

C. Security misconfiguration(a6)

Las páginas web son expuestas a esta vulnerabilidad por su mala configuración de seguridad, estos detalles pueden ocasiones llegar a mostrar mensajes de error que contienen información confidencial, quedando expuestos a que una persona x pueda robar esa información y realizar chantajes a su conveniencia.

1. Protección

Para mejorar la seguridad de las páginas web y estar protegidos ante esta vulnerabilidad se sugiere, configurar de manera idéntica tanto el desarrollo, el control de calidad y los entornos de producción (con diferentes credenciales

utilizadas en cada ambiente). Este proceso debe ser automatizado para minimizar el esfuerzo requerido para configurar un nuevo entorno seguro.

Eliminar o no instalar ninguna característica innecesaria, componentes, documentación y muestras. Eliminar los no usados dependencias y marcos.

D. *Using components with known vulnerabilities*

Provoca ataques que pueden facilitar la pérdida de datos graves o la toma del servidor.

I. *Protección*

Para esta vulnerabilidad se deben eliminar las dependencias no utilizadas, características innecesarias componentes archivos y documentación.

Se debe tener actualizado cada uno de los componentes que van a utilizar, en el lado del cliente y componentes del servidor

Se debe tener en cuenta que toda información personal no siempre se encuentre seguras, aunque existen páginas que se encuentran protegidos ante estas vulnerabilidades, según el análisis realizado existen muchas páginas que por su alta seguridad no permite su escaneo de vulnerabilidades.

V. CONCLUSIONES

Se concluye que la herramienta más apropiada para evaluación de riesgos informáticos de aplicativos web es Acunetix 11, debido a sus excelentes características funcionales, una interfaz amigable, los informes se generan en base a los estándares de buenas prácticas de seguridad, para el caso específico de nuestra investigación el Top Ten de Owasp 2017.

Del análisis realizado en los aplicativos webs de los municipios de la provincia de Cañar, tomando como referencia el Top Ten de Owasp 2017 los datos más relevantes son los siguientes:

El riesgo informático más común dentro de estas páginas web es Broken access control, que consiste en permitir a un usuario acceder a recursos y archivos para los que no debería tener permiso, provocando a los atacantes robar información. Para minimizar este riesgo informático se recomienda realizar configuraciones de cabecera X-Frame-Options en modo DENY que ayuda a que no sea atacado por Clickjacking, además desactivar la cabecera X-Powered-By, lo cual evita que el atacante obtenga información y por último activar la cabecera X-XSS-Protection en modo "1; mode=block" para que el navegador puede detectar ataques XSS y los mismos podrán ser evitados.

Además, otro de los riesgos más frecuentes y que se debe tomar atención es Sensitive data exposure que consiste en que el atacante acceda a los datos confidenciales que no se encuentran protegidos, una de las recomendaciones que se puede plantear es clasificar datos procesados almacenados y transmitidos por un sistema de acuerdo al grado de confidencialidad de los mismos.

Para los riesgos menos frecuentes en el documento que precede se presenta una recomendación para minimizarlos, en vista de que cuando se realiza en forma planificada el aseguramiento de una página web es fundamental tomar en cuenta todos los aspectos relacionados, y sobre todo la prevención es la medida más importante en la seguridad de la información.

REFERENCIAS

- [1] F. N. J. Solarte Solarte, E. R. ENRIQUEZ ROSERO, and M. d. C. Benavides Ruano, "Metodología de análisis y evaluación de riesgos aplicados a la seguridad informática y de información bajo la norma," 2015.
- [2] A. López, *Seguridad informática*. Madrid: EDITEX, 2010.
- [3] C. M. Flores Urgiles and C. H. Flores Urgiles, "ES-PE.edu.ec," 2015.
- [4] "OWASP," 2017.
- [5] j. Mejía Viteri and G. J. Pérez Rueda, "utb.edu.ec," 2017.
- [6] E. G. de Canales Glez, "Generación de reportes de vulnerabilidades y amenazas para aplicaciones web.," 2014.
- [7] J. S. Monar Monar, "http://dspace.esPOCH.edu.ec," 2017.
- [8] J. Orloff, "insureitsecurity.com," 2009.
- [9] G. Gallardo Avilés, *Seguridad en base de datos y aplicativos web*. Madrid: IT Campus Academy, 2015.

Recibido: 3 de mayo de 2018

Aceptado: 15 de junio de 2018



Revisión de algoritmos para la detección de valores atípicos

Review of algorithms for the detection of outliers

Cristina Mariuxi Flores Urgiles^{1*} y Martin Sebastian Ortiz Amoroso¹

Carrera de Ingeniería de Sistemas, Universidad Católica de Cuenca¹

*cmfloresu@ucacue.edu.ec

DOI: https://doi.org/10.26871/killkana_tecnica.v2i1.287

Resumen

La detección de los valores atípicos es una tarea extremadamente importante en una amplia variedad de dominios de aplicación. Con frecuencia estos valores son eliminados para mejorar la precisión de la información, pero a veces la presencia de un valor atípico tiene un cierto sentido o explicación que se puede perder si se elimina, puesto que su identificación puede conducir al descubrimiento de un conocimiento inesperado en diversas áreas como por ejemplo: actividades delictivas en el comercio electrónico, detección de fraudes e incluso el análisis del rendimiento estadístico. El artículo que se presenta es el resultado de una investigación documental, no exhaustiva, de la opinión de diversos autores, que enfocaron su trabajo en determinar la eficiencia de los diversos métodos o algoritmos para la detección de valores atípicos. Inicialmente se realizó un estudio teórico conceptual que permita entender la naturaleza de un valor atípico y su clasificación, para después realizar un análisis sobre las diferentes técnicas de detección basadas en clustering, distancias y densidad. Para cada una de las técnicas de detección de valores atípicos estudiada se presentan algoritmos que han sido planteados por diversos autores además de la eficiencia que cada uno de ellos ha tenido en determinados contextos.

Palabras clave: Valores Atípicos, Minería de Datos, Clustering, Basado en Densidad, Basado en Distancia.

Abstract

The detection of outliers is an extremely important task in a wide variety of application domains. Often these values are eliminated to improve the accuracy of the information, but sometimes the presence of an outlier has a certain sense or explanation that can be lost if it can be eliminated, that its identification can lead to the discovery of unexpected knowledge. Various areas such as: criminal activities in electronic commerce, fraud detection and even statistical performance analysis. The article presented is the result of a non-exhaustive documentary investigation of the opinion of several authors, who focused their work to determine the efficiency of the methods or algorithms for the detection of outliers. Initially, a theoretical conceptual study was carried out to understand the nature of an atypical value and its classification, and then perform an analysis on the different techniques in the determination of clusters, distances and density. For each one of the techniques of detection of atypical values, it was found that the algorithms that have been created by different authors besides the efficiency that each of them has in certain contexts.

Key words: outliers, data mining, Clustering, Density-based, Distance-based.

I. INTRODUCCIÓN

La calidad de los datos que manejan las organizaciones es de gran importancia a la hora de ser analizados para la obtención de información que permita la toma de decisiones empresariales, datos con errores o problemas pueden conducir a obtener información imprecisa y a la vez a tomar decisiones erróneas. Entre los posibles problemas que pueden presentar los datos, se encuentran los conocidos como valores atípicos o "Outliers". Los valores atípicos son el conjunto de objetos que son considerablemente diferentes del resto de los datos[1], considerado además como una observación que se desvía tanto de otras observaciones como para despertar la sospecha de que fue generado por un mecanismo diferente[2]. En muchas de las ocasiones

estos valores son eliminados de tal manera que no afecten los resultados del análisis de los datos, pero en ciertas ocasiones aunque estos valores puedan aparentar que son inválidos y que causarían desviaciones en el análisis de los datos puede ocurrir el sentido contrario. La detección de valores atípicos es un problema extremadamente importante con una aplicación directa en una amplia variedad de los dominios de aplicación, incluida la detección de fraudes[3], identificación de la red informática intrusiones y cuellos de botella[4], actividades delictivas en el comercio electrónico y detección de actividades sospechosas[5], detección de fraude de teléfonos móviles mediante el monitoreo de la actividad telefónica o transacciones sospechosas en los mercados de acciones.

La detección de los llamados Outliers, es una tarea

propia de minería, las personas en la comunidad de minería de datos se interesaron en la detección de valores atípicos después del estudio realizado por Knorr y Ng en donde propusieron un enfoque no paramétrico para la detección de este tipo de valores, que se encuentra basado en la distancia de una instancia a sus vecinos más cercanos[6]. Como se determinó en el estudio[7];[8], no existe un enfoque único o genérico para la detección de datos anómalos, a lo largo del tiempo se han propuesto muchos enfoques para detectar estos tipos de valores, de tal manera que se han clasificado en cuatro categorías según las técnicas utilizadas[9], estas son: basados en distribución, basados en distancia, basados en densidad y basados en clusters o agrupamientos.

Los enfoques basados en la distribución ([2];[10];[11] y [12]) se desarrollan en base a modelos estadísticos a partir de los datos y luego la aplicación de una prueba estadística para determinar si un objeto pertenece a este modelo o no. En el enfoque basado en la distancia abordado por estudios de diferentes autores [[6]; [13]; [14] y [15)], en donde el análisis se establece en base a distancia entre parámetros establecidos por el usuario y los diversos puntos que componen un dataset. Los enfoques basados en la densidad ([16]; [17]) calculan la densidad de las regiones en los datos y declarar los objetos en regiones de baja densidad como valores atípicos. Y por último los enfoques basados en agrupamiento que consideran los grupos de tamaños pequeños como valores atípicos agrupados. En estos enfoques, pequeños clústeres que contienen significativamente menos puntos que otros clústeres son considerados atípicos. [18]

En base a lo mencionado anteriormente en el actual trabajo nos hemos planteado como objetivo realizar un análisis documental de un conjunto de publicaciones de diversos autores, quienes en sus investigaciones han aplicado diversas técnicas para la detección de valores anómalos, con el objeto de determinar la eficiencia de cada uno de ellos; en primer lugar se realiza un análisis del marco teórico sobre la temática a ser analizada, para luego analizar las diferentes técnicas y algoritmos aplicados en cada estudio y sobre todo determinar bajo que contexto fueron utilizados, para finalmente establecer según estos autores que metodologías fueron las más eficientes y que parámetros fueron analizados para llegar a determinarlo.

II. MATERIAL Y MÉTODOS

Para el desarrollo de esta investigación se utilizó como punto de partida un análisis teórico sobre lo que representan los valores atípicos en el tratamiento de la información dentro de la inteligencia de negocios, así también los diversos métodos que existen para la detección de los mismos. Luego, se llevó a cabo el análisis de varios informes de investigaciones relacionadas con respecto a la temática. Análisis realizado con el objetivo de determinar que técnicas y métodos son los más eficientes para la detección de los valores atípicos, concentrándonos en técnicas de minería de datos y de agrupamiento.

A continuación se detalla los aspectos considerados en la metodología utilizada en la presente investigación.

A. Preguntas de investigación

Esta investigación busca responder a las siguientes preguntas de investigación:

- ¿Qué técnicas de minería de datos son consideradas como las más apropiadas para la detección de los datos atípicos?

B. Diseño

Se realizó una revisión sistemática de documentos de sociedades científicas dedicadas a temas de inteligencia de negocios, minería de datos y tratamiento de valores anómalos.

C. Estrategia de búsqueda

En primer lugar se llevó a cabo una búsqueda en Google Scholar de documentos asociados a las siguientes áreas: Técnicas de minería para la detección de datos atípicos, Tratamiento de datos atípicos en la Inteligencia de negocios. Estas búsquedas se realizaron tanto en español como en inglés.

D. Propósito de la Búsqueda

Comprender conceptualmente la naturaleza de un valor atípico y su clasificación. Analizar las diferentes técnicas que se aplican para la detección de un valor atípico. Analizar cada uno de los algoritmos propuestos por diversos autores para la detección de los valores atípicos. Comprender los parámetros utilizados por los autores

E. Fuente de información y Motores de Búsqueda

Artículos científicos, Informes técnicos, Libros. Google Scholar, Bases de datos científicas UCACUE.

F. Criterios de búsqueda

‘Detección de valores atípicos en Minería de datos’, ‘Algoritmos basados en densidad para la detección de valores atípicos’, ‘Algoritmos basados en distancia para la detección de valores atípicos’, ‘Clustering en la detección de valores atípicos’.

G. Criterios de Inclusión

Documentos que contienen información sobre el tratamiento de valores atípicos u outliers.

H. Criterios de Exclusión

Se excluyen los documentos que al referirse a la temática abarcan otro tipo de información que no se relaciona al tema central del análisis, además de excluir aquellos reportes técnicos con datos obsoletos.

I. Evaluación del contenido de los criterios

Exactitud, objetividad, cobertura, relevancia de acuerdo a las preguntas de investigación.

J. Análisis de los datos

La información analizada contribuyó con las siguientes variables: En lo que respecta sobre documentos relacionados con la temática central se pudieron extraer aspectos como fundamentación teórica, conclusiones y resultado de casos de estudio. Los reportes de investigación aportaron con los algoritmos planteados para cada técnica de detección, su funcionamiento y la eficiencia de cada uno de ellos. En base a la información extraída se conformó un documento que proporciona una visión general de la temática, abordando cada aspecto desde el punto de vista del tratamiento de los valores atípicos en búsqueda de un enfoque efectivo para la detección.

III. RESULTADOS Y DISCUSIÓN

A. Valores Atípicos o Outliers

Un valor atípico u Outlier se define como una observación de datos que es muy diferente del resto de los datos observados de una medida específica. A tal punto que a menudo contiene información útil sobre el comportamiento anormal del sistema descrito [18]. Casi todos los estudios que consideran la identificación de valores atípicos como su principal objetivo se encuentran en el campo de las estadísticas y en el de la minería de datos, la detección de los valores atípicos se puede realizar desde una perspectiva univariada o multivariada.

1. Valores Atípicos univariantes

El estudio de los valores atípicos univariantes se centra en el análisis de una única característica o cualidad de un conjunto de datos, por lo cual se supone es fácil conocer cuál es su distribución.

La mayoría de los métodos univariantes más antiguos para la detección de valores atípicos se basan en el supuesto de una distribución subyacente conocida de los datos, que se supone que se distribuye de forma idéntica e independiente. Además, muchas pruebas de discordancia para detectar valores atípicos univariantes suponen además que los parámetros de distribución y el tipo de valores atípicos esperados también son conocidos. Resta decir que, en las aplicaciones de extracción de datos del mundo real, estas suposiciones a menudo se infringen.[19]

Es así que dado un conjunto de datos de n observaciones de una variable x , considerando que \bar{x} es el promedio y s es la desviación estándar de la distribución de datos. Una observación del data set es declarada como un valor atípico si se encuentra fuera del intervalo donde el valor de k esta usualmente tomando como 2 o 3.

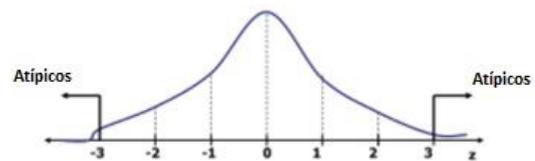
$$(\bar{x} \pm ks) \quad (1)$$

La observación x es considerada como un valor atípico si esta cumple la siguiente condición:

$$\frac{|x - \bar{x}|}{s} > k \quad (2)$$

El problema con los criterios anteriores es que supone una distribución normal de los datos algo que frecuentemente no ocurre. Además, la media y la desviación estándar son altamente sensible a valores atípicos.[20]

Fig. 1. Valores Atípicos univariantes



2. Valores Atípicos Multivariantes

Los valores atípicos multivariantes son observaciones que se consideran extraños no por el valor que toman en una determinada variable, sino en el conjunto de aquellas. Este tipo de valores son mucho más difíciles de detectar que los valores atípicos univariantes, dado que no pueden considerarse “valores extremos”, como sucede cuando se tiene una única variable bajo estudio [21].

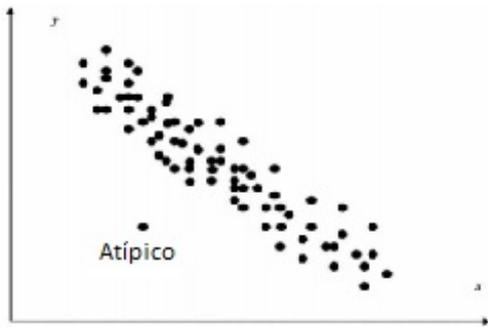
Así también Peat y Barton, manifiestan que un valor atípico multivariante es aquel caso considerado como un valor extremo para una combinación de variables. Por ejemplo, un niño de 8 años de edad cuya estatura sea de 155 cm y pese 45 kg es muy inusual y sería un atípico multivariante[22].

Según Ben-Gal en su estudio denominado “Outlier Detection” manifiesta:

En diversos casos las observaciones multivariantes no pueden ser detectadas como valores atípicos cuando cada variable ha sido considerada de manera independiente[23]. La detección de outliers sólo es posible cuando se realiza un análisis multivariante y las interacciones entre las diferentes variables se comparan dentro de la clase de datos.

El estudio nos presenta un ejemplo que se puede observar en la figura 1, en donde se presenta puntos de datos que tienen dos medidas en un espacio bidimensional. La observación de la parte inferior izquierda es claramente un valor atípico multivariado pero no univariado. Al considerar cada medida por separado con respecto a la difusión de valores a lo largo de los ejes x e y , podemos ver que caen cerca del centro de las distribuciones univariadas. Por lo tanto, la prueba para valores atípicos debe tener en cuenta las relaciones entre las dos variables, que en este caso parecen anormales.

Fig. 2. Valores Atípicos Multivariante

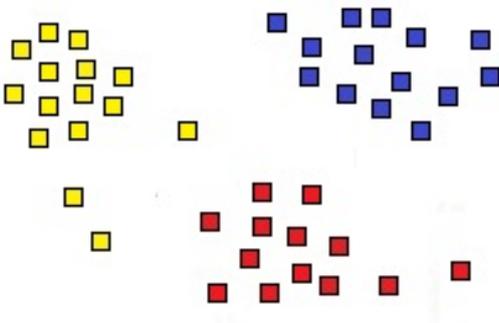


Existen varios métodos para detectar valores atípicos multivariados. Los métodos discutidos en este documento son: detección de valores atípicos por clustering, detección de valores atípicos basados en la distancia y detección de valores atípicos locales basados en la densidad.

B. Detección valores atípicos utilizando clustering

Clustering es una técnica de análisis exploratorio que intenta ordenar los diferentes objetos en grupos, de forma que el grado de asociación entre dos objetos sea máximo si pertenecen al mismo grupo, clustering es una herramienta importante para el análisis de datos atípicos.

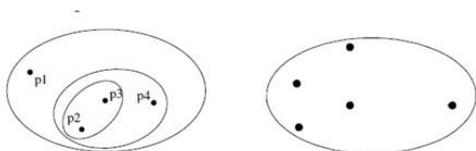
Fig. 3. Clustering



1. Tipos de Clustering

- Clustering Particional: Una división de los objetos de dato en subconjuntos disjuntos (clúster) tal que cada objeto de datos está en exactamente un subconjunto
- Clustering Jerárquico: Un conjunto de clúster anidados organizados como un árbol.

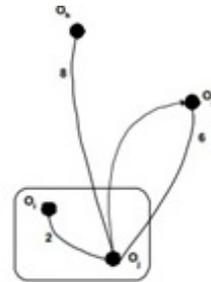
Fig. 4. Clustering



2. Algoritmos de Clustering

- Algoritmo K-Means: En minería de datos k-means es un método de agrupamiento, cuyo objetivo es el de generar k grupos de la partición de un conjunto n de datos, K-means clustering es un método particional en el cual cada clúster está asociado con un centroide (punto central) y en donde cada punto se asigna al clúster con el centroide más próximo, se debe especificar el número de clúster, K.
- Algoritmo PAM (Partitioning Around Medoids): A diferencia del algoritmo Kmeans, el algoritmo de PAM es más robusto y confiable ya que el algoritmo utiliza los objetos más centrales localizados en un clúster (llamados medoide) en lugar de la media del clúster, lo cual resulta más eficiente. PAM trabaja bien en bases de datos pequeñas, pero es lento en grandes.

Fig. 5. Algoritmo PAM, cálculo de distancias entre puntos y medoide



3. Algoritmos de Clustering para la detección de valores atípicos

Se han desarrollado varias técnicas de detección de valores atípicos basadas en clustering. La mayoría de estas técnicas se basan en la suposición clave de que normalmente los objetos pertenecen a grandes y densos clúster, mientras que los valores atípicos muy pequeños. [7][8]

Muchos estudios han abordado la interrogante de si los algoritmos de clustering realmente son técnicas adecuadas para la detección de valores atípicos. Por ejemplo los autores Zhang y Wang informaron en su informe de investigación que los algoritmos basados en clúster no deberían considerarse métodos de detección atípicos.[9] Esto podría ser cierto para algunos de los algoritmos de agrupamiento, como el algoritmo de agrupamiento k-means, esto debido a que los medios de agrupamiento producidos por el algoritmo k-means son sensibles al ruido y a los valores atípicos[24]. El caso es diferente para el algoritmo Partitioning Around Medoids (PAM) ya que este intenta determinar k particiones para n objetos. El algoritmo usa el objeto más céntrico de un clúster llamado medoide en lugar de la media del clúster. PAM es más robusto que el algoritmo k-means en presencia de ruido y valores atípicos. Esto es porque el doids producidos por PAM son representaciones robustas de los centros de clusters y están menos influenciados por valores atípicos y otros valores extremos.[25]

Muchos autores han aprovechado los beneficios que presenta el algoritmo PAM como Moh'd Belal que en su estudio presenta un nuevo método propuesto basado en algoritmos de agrupamiento para la detección de valores atípicos, donde al aplicar el algoritmo PAM considera atípicos a los clúster pequeños y el resto de valores atípicos en el caso de que existieran se calculan en función a las distancias absolutas entre el medoide del clúster actual y cada uno de los puntos en el mismo grupo, los resultados de la prueba mostraron la efectividad del método. Acuna y Rodríguez destacan en su estudio la efectividad del algoritmo CLARA donde se generan múltiples muestras del conjunto de datos, y luego se aplica PAM a la muestra, demostrando en su estudio que PAM es muy robusto ante la presencia de valores atípicos y no depende del orden en que se examinan las instancias.[20]

Por lo contrario Loureiro en su estudio "Outlier Detection Using Clustering Methods: a data cleaning application" describe una metodología de detección de valores atípicos que se basa en clústeres jerárquicos, motivado por la distribución desequilibrada de casos atípicos versus "normales" en estos conjuntos de datos. El autor de mencionado estudio considera a los métodos de clustering parcial como enfoques bastante inestables dado que la mayoría de esos métodos dependen en gran medida de la inicialización de los clusters, y en casi todos los intentos de crear los clusters iniciales, los métodos de agrupamiento no jerárquicos extenderían los valores atípicos en cada uno de ellos. Por lo tanto, son utilizados métodos de agrupación jerárquica, que no dependen de la inicialización de los clusters La idea clave de la propuesta es utilizar el tamaño de los clusters resultantes como indicadores de la presencia de valores atípicos. La suposición básica es que las observaciones atípicas, al ser observaciones con valores inusuales, serán lejanas (en términos de la métrica utilizada para la agrupación) de las observaciones "normales" más frecuentes, y por lo tanto se aislarán en grupos más pequeños[7]

Fig. 6. Algoritmo para identificar valores atípicos en un conjunto de datos, utilizando un método de agrupamiento jerárquico.

Algorithm 1 FindOutliers

```

INPUT:
  DATA, a dataset with  $k$  variables and  $n$  observations;
  a distance function  $d$ ;
  a hierarchical algorithm  $h$ ;
   $nc$  a number of clusters to use (entailing a level of cut of the hierarchy);
  a threshold  $t$  for the size of small clusters.

OUTPUT:
  Out, a set of outlier observations.

Out  $\leftarrow \phi$ 
Obtain the distance matrix  $D$  by applying the distance function  $d$ 
to the observations in  $DATA$ 
Use algorithm  $h$  to grow an hierarchy using the distance matrix  $D$ 
Cut the hierarchy at the level  $l$  that leads to  $nc$  clusters
FOR each resulting cluster  $c$  DO
  IF  $sizeof(c) < t$  THEN
    Out  $\leftarrow Out \cup \{obs \in c\}$ 
    
```

C. Detección de valores atípicos basados en la distancia

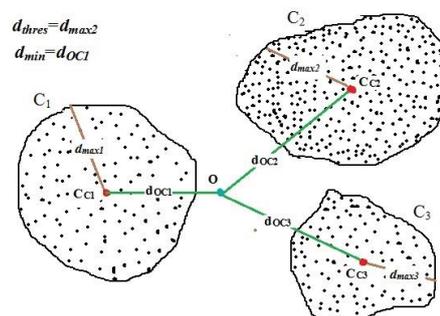
Un método popular para identificar valores atípicos es mediante el examen de la distancia a los vecinos más cerca-

nos de un ejemplo. En este enfoque, uno mira el vecindario local de puntos para un ejemplo típicamente definido por los k ejemplos más cercanos (también conocidos como vecinos). Si los puntos vecinos están relativamente cerca, entonces el ejemplo se considera normal; si los puntos vecinos están muy lejos, entonces el ejemplo se considera inusual. Las ventajas de los valores atípicos basados en la distancia son que no es necesario definir una distribución explícita para determinar la inusualidad, y que se puede aplicar a cualquier espacio de características para el cual podamos definir una medida de distancia.[26]

El método basado en la distancia fue originalmente propuesto por Knorr y Ang[4] donde, un objeto O en un conjunto de datos T es un $DB(p, D)$ -outlier si al menos la fracción p de los objetos en T es mayor que la distancia D desde O . Pero Acuna E. y Rodríguez manifiestan que esta definición tiene ciertas dificultades, como la determinación de D y la falta de una clasificación para los valores atípicos. Por lo tanto, una instancia con muy pocos vecinos dentro de una distancia D puede considerarse como un valor atípico tan fuerte como una instancia con más vecinos dentro de la misma distancia. Además, la complejidad temporal del algoritmo es $O(Kn^2)$, donde k es el número de características y n es el número de instancias. Por lo tanto, no es una definición adecuada para usar con conjuntos de datos que tienen un gran número de instancias.[20]

Por otro lado, Ramaswamy [13], presento en su estudio un algoritmo extendido de detección de valores atípicos basado en la distancia: los n puntos superiores con el D_k máximo se consideran valores atípicos, donde $D_k(p)$ indica la distancia del k -ésimo vecino más cercano de p . Una deficiencia de esta definición es que solo considera la distancia al k -ésimo vecino e ignora la información sobre los puntos más cercanos.

Fig. 7. Enfoque de detección de valores atípicos basado en la distancia. [27]



Varios investigadores han intentado con una variedad de enfoques para encontrar valores atípicos de manera eficiente. Los más simples son los que utilizan bucles anidados [4], [6], [12]. En la versión básica uno compara cada ejemplo con cada otro ejemplo para determinar sus k vecinos más cercanos. Teniendo en cuenta los vecinos para cada ejemplo en el conjunto de datos, simplemente

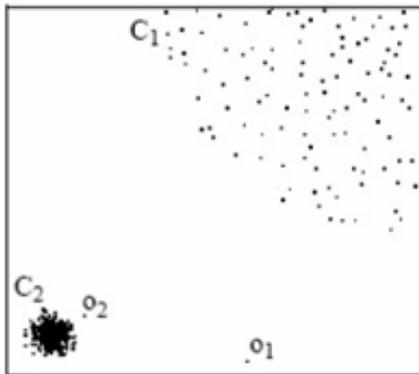
se seleccionan los primeros n candidatos de acuerdo con la definición de valores atípicos. Este enfoque tiene una complejidad cuadrática ya que debemos hacer todos los cálculos de distancia por pares entre los ejemplos.

Bay y Schwabacher presentan un algoritmo para la detección de valores atípicos basado en bucles anidados junto con aleatorización y poda simple que muestra un rendimiento de tiempo casi lineal en muchos data sets conjuntos de datos grandes. Al contrario del rendimiento cuadrático para algoritmos basados únicamente en bucles anidados, la idea principal del algoritmo de bucle anidado es que para cada ejemplo en D se hace un seguimiento de los vecinos más cercanos encontrados hasta el momento. Cuando los vecinos más cercanos de un ejemplo obtienen una puntuación inferior al límite, se elimina el ejemplo porque ya no puede ser un valor atípico. A medida que se procesa más ejemplos, el algoritmo encuentra valores atípicos más extremos y el corte aumenta junto con la eficacia de la poda. Bay usó en su algoritmo la distancia promedio a los k vecinos, pero para Acuna la mediana es más robusta que la media, por lo cual propone una ligera modificación al algoritmo propuesto por Bay.

D. Detección de valores atípicos basados en la densidad

Los enfoques basados en la densidad calculan la densidad de regiones en los datos y declaran los objetos en regiones de baja densidad como valores atípicos. En el estudio "Lof: identifying density-based local outliers", los autores asignan una puntuación atípica a cualquier punto de datos dado, conocido como factor de valor atípico local (LOF), dependiendo de su distancia de su vecindario local[27]. Se informa un trabajo similar del autor Papadimitriou[28]. La siguiente figura tomada del estudio realizado por Breuning muestra la debilidad del método basado en la distancia que identifica como atípica la instancia O1, pero no considera como un valor atípico O2.

Fig. 8. Ejemplo para mostrar la debilidad del método basado en la distancia para detectar valores atípicos. [27]



Por otro lado Papadimitriou en su estudio propone un nuevo método para evaluar valores atípicos, llamado la

integral de correlación local (LOCI). Al igual que con los mejores métodos anteriores, LOCI es muy eficaz para detectar valores atípicos y grupos de valores atípicos. Además, proporciona un punto de corte automático basado en razonamiento probabilístico, dictado por datos para determinar si un punto es un valor atípico; en contraste, los métodos previos obligan a los usuarios a elegir puntos de corte, sin ningún indicio de que qué valor de corte es mejor para un conjunto de datos determinado. Además incluye el factor de desviación de granularidad múltiple (MDEF), que puede hacer frente a las variaciones de densidad local en el espacio de características y detectar tanto los valores atípicos aislados como los clústeres periféricos. [28]

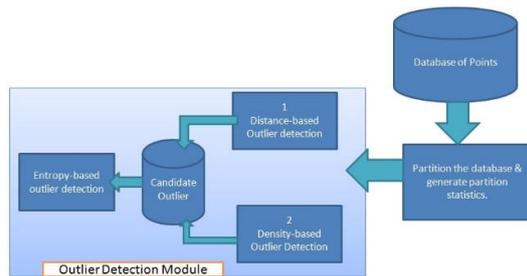
E. Detección de valores atípicos basados en la distancia y densidad

El estudio realizado por A. Mira y sus colegas denominado "RODHA: Robust Outlier Detection using Hybrid Approach" presenta el desarrollado un robusto algoritmo de detección de valores atípicos supervisados utilizando el enfoque híbrido (RODHA) que incorpora tanto el concepto de distancia y densidad junto con la medida de entropía como la determinación de un valor atípico.

El estudio realizado por A. Mira y sus colegas denominado "RODHA: Robust Outlier Detection using Hybrid Approach" presenta el desarrollado un robusto algoritmo de detección de valores atípicos supervisados utilizando el enfoque híbrido (RODHA) que incorpora tanto el concepto de distancia y densidad junto con la medida de entropía como la determinación de un valor atípico. La efectividad del algoritmo resulta de los enfoques combinados de detección de valores atípicos basados en la distancia y la densidad. El enfoque basado en la distancia por sí solo es capaz de detectar valores atípicos para conjuntos de datos donde los objetos se distribuyen uniformemente entre los grupos de datos. La debilidad del enfoque basado en la distancia para detectar valores atípicos para conjuntos de datos distribuidos de manera no uniforme se compensa por el enfoque basado en la densidad al considerar la densidad local en torno a un objeto atípico candidato.

Además, la incorporación de entropía para la detección de valores atípicos lo hace más robusto y sensible que otras técnicas existentes de detección de valores atípicos. El cálculo de la entropía dentro del clúster utilizando la medida de entropía tiene una ventaja, ya que se presta muy bien a la estimación no paramétrica directamente de los datos[29] y considera cómo se distribuyen los datos dentro del clúster. El algoritmo RODHA propuesto tiene una complejidad de tiempo lineal. El rendimiento de detección del algoritmo es competitivo excelente que otros algoritmos existentes.[30]

Fig. 9. Ejemplo para mostrar la debilidad del método basado en la distancia para detectar valores atípicos.[27]



IV. CONCLUSIONES

En el presente documento, se ha realizado un análisis de varios informes de investigación sobre la temática abordada, en donde se ha podido estudiar varios métodos planteados por diversos autores para la detección de los valores atípicos. Cada uno de los autores ha realizado un aporte importante ya que han presentado algoritmos que han sido acoplados a las necesidades de cada estudio, para cada estudio se ha determinado la eficiencia del algoritmo en términos de la complejidad. De los informes de investigación analizados se presenta a continuación los aportes más relevantes:

- Según Zhang y Wang los algoritmos basados en clúster no deberían considerarse métodos de detección atípicos, de manera específica se hace referencia al algoritmo k-means, esto debido a que los medios de agrupamiento producidos por el algoritmo k-medias son sensibles al ruido y a los valores atípicos.
- Caso contrario es el algoritmo de PAM, que ha sido utilizado por varios autores para la detección de valores atípicos considerando a este como más robusto ante la presencia de valores atípicos.
- Aunque existen autores con opiniones contrarias en donde consideran a este enfoque inestable ya que extenderían los valores atípicos a cada clúster y en lugar de este utilizan algoritmos jerárquicos considerados como más efectivos.
- Los algoritmos basados en distancia propuestos por Bay y Acuna demuestran mayor eficiencia en términos de complejidad, ya que cuentan con un rendimiento de tiempo casi lineal. Al contrario del rendimiento cuadrático para algoritmos basados únicamente en bucles anidados.
- Los algoritmos basados en densidad presentan ventajas sobre los algoritmos basados en distancia, en su estudio Breunig muestra la debilidad del enfoque basado en la distancia para detectar valores atípicos en conjuntos de datos distribuidos de manera no uniforme.
- Para solventar los inconvenientes propuestos se ha creado un algoritmo con un enfoque híbrido (RODHA) que incorpora tanto el concepto de distancia y densidad junto con la medida de entropía como la determinación de un valor atípico. En donde la debilidad del enfoque

basado en distancia se compensa por el enfoque basado en la densidad al considerar la densidad local en torno a un objeto atípico candidato.

REFERENCIAS

- [1] J. Han and M. Kamber, "Data Mining: Concepts and Techniques,"
- [2] D. M. Hawkins, *Identification of Outliers*. London: Chapman & Hall, 1980.
- [3] R. a. D. J. H. Bolton, "Statistical Fraud Detection: A Review," *Statistical Science*, pp. pp. 235–249, 2002.
- [4] T. a. C. E. B. Lane, "Temporal Sequence Learning and Data Reduction for Anomaly Detection," *ACM Transactions on Information and System Security*, pp. Pages 295–331, 2000.
- [5] A. a. A. F. Chiu, "Enhancement on Local Outlier Detection.," *Chiu, A. an 7th International Database Engineering and Application Symposium (IDEAS03)*, pp. pp. 298–307., 2003.
- [6] E. a. R. N. Knorr, "Algorithms for Mining Distance-based Outliers in Large Data Sets," *Proc. the 24 th International Conference on Very Large Databases (VLDB)*, pp. pp. 392–403., 2000.
- [7] A. Loureiro, "Outlier Detection using Clustering Methods: a Data Cleaning Application," in *Proceedings of KNet Symposium on Knowledge-based Systems for the Public Sector, Bonn, Germany*.
- [8] K. Niu, "ODDC: Outlier Detection Using Distance Distribution Clustering," *PAKDD 2007 Workshops, Lecture Notes in Artificial Intelligence (LNAI) 4819, Springer-Verlag.*, pp. pp. 332–343, 2007.
- [9] J. a. H. W. Zhang, "Detecting outlying subspaces for high-dimensional data: the new Task, Algorithms, and Performance," *Knowledge and Information Systems.*, 2006.
- [10] V. a. T. L. Barnett, "Outliers in Statistical Data," *John Wiley.*, 1994.
- [11] P. a. A. L. Rousseeuw, *Robust Regression and Outlier Detection*. John Wiley & Sons., 2000.
- [12] E. Knorr, "Distance-based Outliers: Algorithms and Applications.," *VLDB Journal*, pp. 237–253., 2000.
- [13] S. Ramaswami, "Efficient Algorithm for Mining Outliers from Large Data Sets," *Proc. ACM SIGMOD*, pp. pp. 427–438., 2000.
- [14] F. a. C. P. Angiulli, "Outlier Mining in Large High-Dimensional Data Sets," *IEEE Transactions on Knowledge and Data Engineering*, 17(2), pp. 203–215, 2005.
- [15] H. K. M., "Lof: identifying density-based local outliers," In *Proceedings of 2000 ACM SIGMOD International Conference on Management of Data.*, pp. 93–104, 2000.
- [16] H. K. S., "Fast outlier detection using the local correlation integral.," *Proc. of the International Conference on Data Engineering*, pp. pp. 315–326., 2003.

- [17] J. Almeida, "Improving Hierarchical Cluster Analysis: A New Method with Outlier Detection and Automatic Clustering," *Chemometrics and Intelligent Laboratory Systems*, pp. 208–217, 2007.
- [18] C. C. Aggarwal, "An effective and efficient algorithm for high-dimensional outlier detection," *The VLDB Journal*, vol. 14, pp. 211–22, 2005.
- [19] V. a. L. Barnett, *Outliers in Statistical Data*. John Wiley., 2000.
- [20] A. E. a. R. C., "A Meta Analysis Study of Outlier Detection Methods in Classification, Technical paper, Department of Mathematics, University of Puerto Rico at Mayaguez," tech. rep., 2004.
- [21] J. R. K. R. Gnanadesikan, "Robust Estimates Residuals and Outlier Detection with Multiresponse Data," *Biometrics.*, vol. 28, pp. pp 81–124.
- [22] B. B. J. Peat, "Medical Statistics: "A guide to data analysis and critical appraisal"," *Blackwell Publishing*, 2005.
- [23] I. Ben-Gal, "Outlier detection," *Data Mining and Knowledge Discovery Handbook: A Complete Guide for Practitioners and Researchers*, pp. 131–146, 2005.
- [24] A. a. R. D. Jain, *Algorithms for Clustering Dat.* Prentice-Hall., 1988.
- [25] M. Laan, "A New Partitioning Around Medoids Algorithms," *Journal of Statistical Computation and Simulation*, 2003.
- [26] S. D. Bay, *Mining distance-based outliers in near linear time with randomization and a simple pruning rule.* 2003.
- [27] M. Breunig, "identifying density-based local outliers," *Proceedings of 2000 ACM SIGMOD International Conference on Management of Data*, pp. 93–104., 2000.
- [28] S. Papadimitriou, "Fast outlier detection using the local correlation integral.," *Proc. of the International Conference on Data Engineering*, pp. pp. 315–326., 2003.
- [29] D. X. J. Principe, "Unsupervised Adaptive Filtering," in *Information Theoretic Learning*, vol. 1, John Wiley & Sons, 2000.
- [30] D. K. B. A. Mira, "RODHA: Robust Outlier Detection using Hybrid Approach," *American Journal of Intelligent Systems*, 2012.

Recibido: 3 de mayo de 2018

Aceptado: 15 de junio de 2018



Modelos de Referencia de Arquitectura Empresarial para la Industria de Educación Superior

Enterprise Architecture Frameworks for Higher Education Industry

Blanca Lucia Avila Correa^{1*}

Unidad Académica de Tecnologías de Información y Comunicación, Universidad Católica de Cuenca

*blavilac@ucacue.edu.ec

DOI: https://doi.org/10.26871/killkana_tecnica.v2i1.288

Resumen

La digitalización de las universidades es uno de los grandes desafíos de la universidad ecuatoriana en la actualidad. La búsqueda de la calidad en la educación superior implica el logro de transformaciones desde la perspectiva estratégica, por lo que los líderes de TI deben estar preparados para formar parte del equipo de toma de decisiones que entienda las necesidades empresariales, proponga soluciones tecnológicas que garanticen la generación de valor. A lo largo del tiempo, la falta de entendimiento de las necesidades de los grupos de interés, ha llevado a proponer soluciones aisladas y desconectadas, que han abordado los problemas sin hacer un análisis causal. El presente trabajo, busca motivar a los líderes de TI de las instituciones de educación superior del Ecuador la implementación de procesos de arquitectura empresarial a plantear soluciones estratégicas y aborden la problemática de las universidades de forma integral, articulando los componentes estratégicos, del negocio, de las aplicaciones, datos y la tecnología a través de arquitectura empresarial. La propuesta se basa en el marco de referencia TOGAF y su metodología ADM, a la que se suman una lista de marcos de referencia, modelos, estándares y buenas prácticas alineadas a la educación superior. Estos referentes deberán ser adoptados en función del contexto la necesidad de cada institución. en el entorno en el que se aplican. Cabe mencionar que en este trabajo se han omitido otras herramientas y métodos propios del marco de referencia, los cuales deberán ser adoptados en el ejercicio de la arquitectura empresarial.

Palabras clave: Área de la Organización, Dominio, Iniciativa, Problema, Marco de referencia.

Abstract

The digitalization of the universities is one of the great challenges of the Ecuadorian university at present. The search of the quality in the higher education implies the achievement of transformations from the strategic perspective. For that reason, the CIOs must be prepared to take part to the decision-making team, which understands the business needs, proposes technological solutions that guarantee the generation of value. Over time, the lack of understanding of the needs of stakeholders, has led them to propose isolated and disconnected solutions, which have addressed the problems without doing a causal analysis. The present work seeks to motivate to CIOs of the higher education institutions of Ecuador to implement business architecture to propose strategic solutions and address the problems of the universities in an integral manner, articulating the strategic components of the business, Applications, data and technology through business architecture. The proposal from the TOGAF reference framework and its ADM methodology is add a list of reference frameworks, models, standards and good practices aligned to higher education. These referents should be adopt depending on the context the need of each institution in the environment in which they are applied. It is worth mentioning that in this work other tools and methods of the frame of reference have been omit, which should be adopt in the exercise of the business architecture.

Key words: Enterprise, Domain, Initiative, Concern, Framework.

I. INTRODUCCIÓN

El relevante papel de las TICs en las organizaciones, exige que los departamentos de TI propongan soluciones tecnológicas que logren una verdadera transformación estratégica.

La consolidación de las organizaciones, requiere solu-

ciones estratégicas de tecnología, con equipos gerenciales que tengan un alto nivel de liderazgo; entonces, el líder de TI, debe formar parte del equipo de toma de decisiones y lograr resultados que generen valor al negocio y resuelvan problemas empresariales. Desde esta perspectiva, el líder de TI debe ser un estratega; que entienda las necesidades

empresariales, proponga soluciones y plantee alternativas de logro.[1]

Puesto que la Arquitectura Empresarial es la disciplina encargada de entregar soluciones a los problemas empresariales; debe satisfacer las necesidades empresariales de los grupos de interés (accionistas, empleados, clientes, proveedores, entidades regulatorias, etc.). Entendiéndose por necesidades empresariales al incremento de la rentabilidad, de la satisfacción de los clientes, la reducción de costos o la optimización de las inversiones.

Entonces, ¿Cómo puede TI entender todas estas necesidades, traducirlas en procesos de negocio, mantener el modelo jerárquico dentro de la empresa y buscar que estos procesos tengan soluciones tecnológicas, innovadoras, sostenibles, estructuradas e integrales?; la respuesta está en la implementación de Arquitectura Empresarial.

Ahora, ¿Cómo reconocemos que una organización requiere implementar trabajos de arquitectura empresarial? Debido al incesante cambio que sufren muchas empresas, su adaptación al entorno dinámico requiere desarrollar un sentido de urgencia, que contrarreste con esta realidad. Esto frecuentemente termina haciendo que TI proponga soluciones tecnológicas aisladas y temporales. Por otro lado, la complejidad y la premura convergen en una necesidad de priorización y jerarquización de los problemas estratégicos, tácticos y operativos. Su desarticulación, produce soluciones aisladas y el desconocimiento de la problemática por parte de los niveles operativos, haciendo que la organización se desgaste en soluciones que no agregan valor.[1]

Para ilustrar cómo las organizaciones pueden consolidarse a través de los procesos de Arquitectura Empresarial, el presente trabajo pretende orientar al lector sobre el proceso de arquitectura empresarial con enfoque en la educación superior. La elección de este sector se hace en función de los innumerables cambios a los que se han sometido las instituciones de educación superior en el Ecuador desde el año 2010, luego de que entrara en vigencia la *Ley de Educación Superior* (LOES), además de la dinámica de las tendencias tecnológicas cuya adopción las vuelve altamente competitivas y sostenibles. En este contexto, el trabajo se enfocará en puntualizar el proceso y los referentes más relevantes para el desarrollo de una arquitectura empresarial en las universidades.

II. FUNDAMENTOS DE LA ARQUITECTURA EMPRESARIAL

La arquitectura empresarial, se ha consolidado como una propuesta de solución de problemas empresariales que abordan la raíz de los problemas de manera estructurada para encontrar soluciones definitivas apoyadas en tecnología.

En muchas organizaciones, el entendimiento a la problemática empresarial es sintomático; es decir, no se hace un análisis causal antes de definir un esquema de solución. La arquitectura empresarial fue creada para que los responsables de TI puedan erradicar los problemas haciendo una interpretación de la realidad a través del análisis causal y

proponiendo soluciones preventivas, proactivas y predictivas. [1]

En el caso de las universidades del Ecuador, los factores de sostenibilidad y competitividad, crean la necesidad urgente de apalancamiento de los procesos académicos, de investigación y de vinculación con la sociedad; con soluciones tecnológicas estratégicas que reduzcan la abrumadora brecha de transformación digital que tienen respecto las exigencias del mercado internacional. Las tendencias tecnológicas están propuestas en los informes del Espacio Europeo de Educación Superior (EEES), el informe *Horizon del New Media Consotium*, los informes anuales de EDUCAUSE y Gartner entre otros.[2]

A. Dominios de la Arquitectura Empresarial

Cuando hacemos un visionamiento del ecosistema institucional de las universidades, al igual que en otras organizaciones se evidencian componentes estratégicos, de negocio y tecnológicos. Precisamente, desde esta perspectiva, la arquitectura empresarial enfoca el tratamiento de las soluciones empresariales mediante la articulación de las estrategias, los componentes de negocio (procesos y organización), la información, las aplicaciones, la infraestructura tecnológica y los riesgos y seguridades.[3]

Todos los marcos de trabajo de la arquitectura empresarial, abordan de una u otra forma estos componentes a los que se denominan *Dominios de la Arquitectura Empresarial* y su clasificación puede variar. Por ejemplo; para el padre de la arquitectura empresarial, John Zachman; esta disciplina se consolida en cinco dominios: Estrategias y Normatividad, Estructura, Procesos de Negocio, Información y Servicios de TICs; mientras que en el marco de referencia TOGAF (*The Open Group Architecture Framework*), éstos se agrupan en cuatro dominios: Negocio, Aplicaciones, Información y Tecnología. Este último enfoque, será tratado en el presente trabajo.[2]

De cualquier modo, la articulación de estos dominios, establece un lenguaje común de entendimiento, para que cada uno genere valor a la organización.

1. Dominio del Negocio

En este dominio se definen los componentes estratégicos y del negocio. El análisis de los componentes estratégicos comprende: los grupos de interés, los elementos de direccionamiento estratégico (misión, visión, políticas, diagnóstico situacional, mapa estratégico y entorno organizacional) el análisis competitivo, los servicios del negocio y el modelo de negocio.[4]

En el análisis del negocio se deben articular; la cadena de valor, los procesos estratégicos de apoyo; así como la organización, sus roles y competencias.[5]

2. Dominio de Aplicaciones

Consiste en el análisis de las plataformas de *Core*, las aplicaciones existentes y requeridas por la organización. De cada aplicación debe especificarse un plano de los sistemas

y sus interrelaciones con los procesos más importantes y la información que gestionan.[5]

Es importante que este dominio logre que la organización alinee sus aplicaciones con las necesidades del negocio. Para ello se deberá identificar cuáles son las brechas entre la arquitectura de aplicaciones actual y la propuesta, a través de una evaluación de las capacidades, un plan de comunicación y el modelo organizacional.

3. Dominio de Información

Consiste en la descripción de los datos asociados a con las aplicaciones. Es decir, los datos de los macro-procesos que son parte de los datos corporativos y que apoyan la gestión a través de bases de datos, sistemas de gestión de bases de datos, repositorios de información, etc. Adicionalmente se analizarán los métodos de administración, los flujos de datos, los modelos de análisis de desempeño y toma de decisiones, los maestros de datos y las políticas de seguridad de la información.[6]

Es necesario que, en este dominio se consideren aspectos anclados a los sistemas de aplicación que los procesan, cómo se relacionan entre sí y la relación de los sistemas con los procesos que los gobiernan.

4. Dominio de Tecnología

En este dominio se debe describir la situación tecnológica actual respecto al software de base, la infraestructura física, las redes y telecomunicaciones; así como los componentes de seguridad requeridos.[6]

El análisis debe considerar cuáles son las necesidades de nuevas soluciones tecnológicas. Cada uno de los componentes debe contener una arquitectura actual y propuesta que se derive de las brechas encontradas para ser analizadas.[2]

B. Proceso de Arquitectura Empresarial para las Instituciones de Educación Superior

El proceso de arquitectura empresarial se basa en la taxonomía de los componentes de los dominios y la ontología entre los objetivos estratégicos de negocio con los componentes tecnológicos de cada dominio.[7]

Para abordar la complejidad de las organizaciones mediante la articulación de los dominios; es necesario guiar el ejercicio de arquitectura empresarial con un marco de trabajo que establezca los lineamientos, metodología y técnicas.

Los marcos de referencia para el desarrollo de la arquitectura empresarial más comunes son: Zachman, FEA (*Federal Enterprise Architecture*), Gartner y TOGAF.[8]

El marco de referencia de Zachman concibe el desarrollo de la arquitectura empresarial como una desagregación de los componentes empresariales tales como: estrategias, procesos, organización, información, aplicaciones e infraestructura. Esta organización permite responder al qué, cómo, dónde, quién, cuándo y el porqué de un proyecto en cada

dominio. El proceso taxonómico define hasta 36 subniveles de estratificación de los dominios. [8]

El marco de referencia FEA (*Federal Enterprise Architecture*), propuesto por el gobierno de los Estados Unidos, está centrado en describir y analizar las inversiones de TI, mejorar la colaboración y transformar el gobierno para el desarrollo de la arquitectura empresarial. Éste marco de referencia relaciona el rendimiento, negocio, servicios, datos para establecer la alineación estratégica entre el negocio y la tecnología.[9]

El marco de referencia de Gartner se enfoca en el incremento de la rentabilidad a través de la agrupación de los propietarios del negocio, los especialistas de la información y los implementadores de tecnología. Este modelo utiliza la metodología GEAM (*Gartner Enterprise Architecture Method*) para fusionar la arquitectura del negocio, la arquitectura de las aplicaciones y datos y la arquitectura de tecnologías en el contexto organizacional.[10]

La propuesta del Open Group, para el desarrollo de la Arquitectura Empresarial es TOGAF. Es considerado por los especialistas como un referente agnóstico y está disponible de forma libre en la web. [2] Estas características favorecen la implementación de la arquitectura empresarial en las instituciones de educación superior del Ecuador, por lo que será explicado con mayor detalle.

TOGAF fundamenta el trabajo de arquitectura empresarial en los dominios del negocio, información, aplicaciones y tecnología. Se basa en el ADM (*Architecture Development Method*), el cual describe el procedimiento que va desde el entendimiento de una necesidad de Arquitectura Empresarial hasta la solución del problema empresarial (*concern*) [11]. El ADM de TOGAF se define en diez fases (figura 1.), cada una de las cuales provee al arquitecto empresarial; la guía de desarrollo de la arquitectura empresarial, los entregables, artefactos y la validación de los resultados. Este método es iterativo en cada proceso de arquitectura empresarial y entre sus fases lo que permite avanzar de forma sistemática e incremental o incluso volver al punto de partida.

1. Fase Preliminar

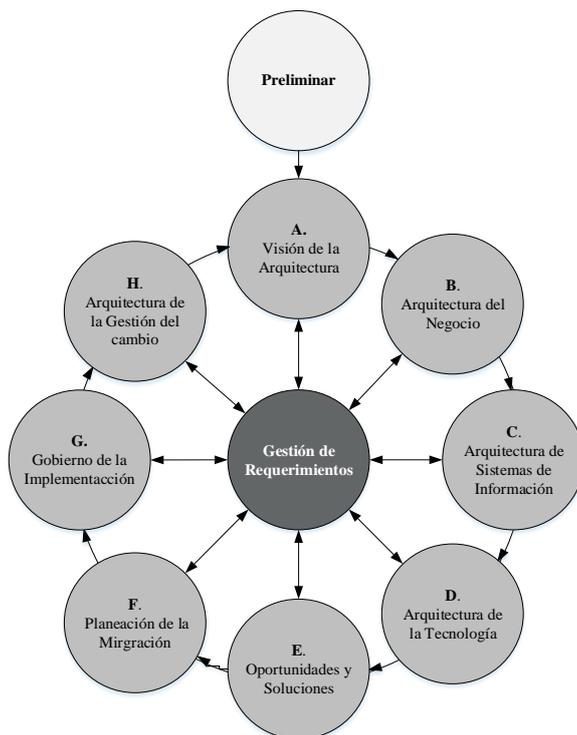
La *fase preliminar* es la etapa de alistamiento de la arquitectura empresarial, en ésta se resuelve qué tan preparada está la organización para acoger un ejercicio de arquitectura empresarial. Es decir, cuál es la capacidad de la empresa para realizar la arquitectura empresarial. Para ello se debe determinar; dónde, qué, porqué, quién y cómo se va a desarrollar la arquitectura empresarial[11]. Esto significa definir el contexto o área de la organización a considerar (*enterprise*), cuál es el problema (*concern*), los principios o reglas a corto plazo que regirán el proceso, el gobierno y el equipo de trabajo que formulará la arquitectura; además de las técnicas y guías que apoyarán el proceso.

Quienes lideran el proyecto, deberán estructurar y desarrollar todos los elementos que se relacionen con los antecedentes y el contexto de la empresa, la metodología y la toma

de decisiones informadas sobre la aplicación del marco de trabajo.

La identificación de los grupos de interés y sus necesidades, un plan de riesgos, el modelo de negocio que identifique; los objetivos estratégicos, la cultura organizacional y jerarquía en la toma de decisiones; los procesos actuales documentados, las metodologías de gestión, el portafolio de proyectos, el mapa de ruta con el presupuesto alineado entre otros; deben estar disponibles antes de arrancar el proceso de arquitectura empresarial.

Fig. 1. Esquema de la Metodología ADM de TOGAF



Fuente: The Open Group

2. Fase A: Visión de la Arquitectura

La fase visionamiento de la arquitectura, consiste en definir un escenario de llegada de la arquitectura; es decir, cuáles son las capacidades que se quieren desarrollar. Se trata de obtener una vista panorámica a través de un reconocimiento de alto nivel que incluya todos los componentes estratégicos y de dirección de la institución que respondan a los motivadores de negocio, nuevos escenarios y cambios de normativas.

En esta fase el equipo de arquitectura debe considerar evaluar la misión, visión, valores, gobierno, estrategias tanto del negocio como de TI; analizar la situación actual mediante FODA (Fortalezas, Oportunidades, Debilidades y Amenazas) o ERIC (Eliminar, Reducir, Incrementar o Crear); analizar la competencia, la competitividad, los grupos de interés, el portafolio de productos y servicios; definir el modelo de negocio y los requerimientos funcio-

nales y no funcionales que apoyan la formulación de la arquitectura[11].

3. Fase B: Arquitectura del Negocio

En esta fase corresponde definir una línea de partida, línea base o *Base Line* de los componentes del negocio. Ello implica definir los procesos, servicios y funciones.

En este punto, el equipo de arquitectura empresarial debe determinar los motivadores o razones que apalancan el desarrollo de la arquitectura, analizar la organización (organigrama institucional), funciones, personas y roles institucionales, así como sus inter-relaciones. Analizar el mapa de procesos institucional con el detalle de los procesos estratégicos y misionales (académicos, de investigación y de vinculación con la sociedad) y los de soporte.[2]

Asimismo, será necesario determinar los productos o servicios, como la oferta académica de pregrado, posgrado, las diferentes modalidades de estudio entre otros.[2]

4. Fase C: Arquitectura de los Sistemas de Información

En este punto del proceso, el equipo deberá usar esquemas de arquitectura que definan la línea base de las aplicaciones clave que gestionan los procesos institucionales, tales como plataformas de Core, de servicios, interfaces y canales, aplicaciones móviles, aplicaciones para la toma de decisiones, plataformas de integración y bus de servicios, etc.[12] Además se debe establecer su relación con los procesos para visualizar necesidades de implementación.

De igual manera, el equipo de arquitectura empresarial, deberá definir esquemas de arquitectura que identifiquen los activos de información institucionales como; bases de datos relacionales, no relacionales, su administración y consistencia, los repositorios de contenidos digitales, modelos de análisis de desempeño de inteligencia de negocios y toma de decisiones, datos maestros institucionales, sistemas de gestión de seguridad informática. La profundidad del análisis deberá contemplar las relaciones de éstos con las aplicaciones que los gestionan.[2]

5. Fase D: Arquitectura de la Tecnología

La definición de la línea base de la infraestructura tecnológica deberá hacerse especificando los componentes tecnológicos y los equipos de seguridad, la infraestructura de redes y comunicaciones de amplia cobertura, las redes locales e inalámbricas, la infraestructura de almacenamiento y los centros de datos, los clústeres de virtualización, y el software de bases de datos. En este último se deberán incluir el licenciamiento, niveles de estandarización y principalmente las políticas de seguridad.[2]

En cada uno de los modelos arquitectónicos de línea base definidos en las fases, B, C, y D; deberá existir una arquitectura propuesta de solución que determina hacia dónde se quiere llegar; es decir, la arquitectura objetivo o target. Esta propuesta debe estar basada en las tendencias, estándares, modelos, marcos de referencia y mejores prácticas propias de la industria de la educación superior. Este escenario,

debe asegurar el cumplimiento de los requerimientos de los grupos de interés y los problemas identificados.[12]

Sobre esta base, lo siguiente es establecer las diferencias entre el escenario actual y el futuro (*baseline* y *target*) para definir y analizar las brechas de cada uno de los elementos de los dominios de la arquitectura empresarial. Una de las herramientas más usadas en este punto es el *Modelo de Madurez de las Capacidades* que permite evaluar cada componente en una escala que va de 0 (incompleto) a 5 (optimizado).[2] En función del resultado de la evaluación se define el escenario *target*, el cual representa el referente más próximo para lograr un cambio significativo.

Del *GAP Analysis* o análisis de brechas, se desprenden las iniciativas de solución que pasarán a formar parte de una hoja de ruta o *Roadmap*. [11]

6. Fase E: Oportunidades y Soluciones

Las brechas detectadas representan las necesidades, insuficiencias o carencias que tienen la institución; es decir las iniciativas. La agrupación lógica de iniciativas conformará los paquetes de trabajo de arquitectura empresarial. Aquí, se afinan detalles del Mapa de Ruta de manera que se convierta en una herramienta estratégica de conducción de los paquetes de arquitectura empresarial hacia la consecución de las necesidades de los grupos de interés. Las soluciones y oportunidades definidas, deberán también contemplar las limitaciones de ejecución. [11]

Las iniciativas empaquetadas, deberán seleccionarse en base a los motivadores del negocio, los cuales generalmente buscan incrementar la rentabilidad, la satisfacción de los clientes, reducir los costos u optimizar las inversiones.

La versión completa del mapa de ruta, deberá incorporar los costos, el tiempo, los recursos y un análisis de las implicaciones que tendrá la implementación de las arquitecturas. Este mapa de ruta representa el plano de la capacidad de transición de la arquitectura y deberá estar acompañado de un plan de implementación y migración.

7. Fase F: Planeación de la Migración.

La fase anterior, define la capacidad de la institución para la transición de la línea base a la arquitectura objeto, considerando la relación costo-riesgo-beneficio. En este punto, el equipo de arquitectura empresarial, debe detallar el plan de implementación y migración para hacer realidad la arquitectura *target* definiendo entregables, hitos y listas de verificación que permitan hacer un seguimiento y monitoreo del proceso. Este plan debe estar coordinado con los procesos de gestión del portafolio y de proyectos de la institución. Asimismo, deberá considerar aspectos de gestión del cambio que incluyan actividades de evaluación de dependencias, costos y beneficios de cada proyecto en el contexto institucional. [11]

8. Fase G: Gobierno de la Implementación

El Plan de Migración e Implementación, se elabora por la necesidad de supervisar la implementación de acuerdo a

lo establecido en él. El cumplimiento de este plan requiere de la definición de Acuerdos de Niveles de Servicio (SLA) y una adecuada Gestión de Proveedores. [2]

Para reducir los riesgos del fracaso en la implementación de los paquetes de arquitectura, deberá supervisarse exhaustivamente. De esto se encarga la función de Gobierno de la Arquitectura. El aseguramiento de la implementación, implica que ésta se realice con una serie de transiciones incrementales hacia la arquitectura deseada. [11]

9. Fase H: Gestión de Cambio de la Arquitectura

El objetivo de esta fase es cumplir con el proceso de mejora continua a través de la evaluación constante del trabajo de arquitectura, que detecte resultados desfavorables en los ciclos de implementación de la arquitectura.[1]

Cada ejercicio de arquitectura, deberá estar acompañado de una adecuada gestión del cambio que garantice el cumplimiento de la arquitectura deseada que apunta a agregar valor al negocio. En este punto es necesario que el equipo de arquitectura considere alternativas de *roll back* que permitan regresar a un estado anterior de ser necesario. [11]

La gestión del cambio, se encargará de supervisar continuamente nuevas solicitudes, desarrollos tecnológicos o cambios del entorno empresarial. Una vez identificado el cambio, el proceso de gestión de cambio determinará si es necesario iniciar formalmente un nuevo ciclo evolutivo de la arquitectura y permitirá establecer criterios para juzgar si una solicitud de cambio requiere actualización o necesita un nuevo ciclo.

III. MARCOS DE REFERENCIA COMPLEMENTARIOS PARA EL TRABAJO DE LA ARQUITECTURA EMPRESARIAL EN LA EDUCACIÓN SUPERIOR

Los marcos de referencia de la arquitectura empresarial, guían el trabajo para llevar a las organizaciones a un escenario deseado. Este trabajo se ejecuta con el apoyo de estándares, modelos, marcos de referencia y buenas prácticas propios de cada vertical de negocio.

Las instituciones de educación superior, por naturaleza, tienen componentes, organizaciones y procesos complejos que han llevado a varios organismos y firmas internacionales a proponer metodologías, técnicas y herramientas que pueden ser aplicados durante el proceso de arquitectura empresarial.

BMM (Business Motivation Model), usado con frecuencia en los ejercicios de arquitectura empresarial para justificar las razones por las que se debe realizar el trabajo. Permite la planeación de las iniciativas y la evaluación rigurosa de los resultados. [2]

El Análisis de las Cinco fuerzas Competitivas de Porter, es un referente que apoya el análisis de la competitividad. Las universidades pueden hacer uso de esta herramienta para definir su posición frente a la competencia y hacer frente a ella asegurando su sostenibilidad.[13]

La Matriz BCG (*Boston Consulting Group*), permite a las universidades autofinanciadas definir un portafolio de productos o servicios que generen mayor valor. BCG optimiza la oferta académica, a través de un reconocimiento de los niveles de crecimiento y participación del mercado. Aquellas que tienen bajos niveles de participación o crecimiento, representan un problema para la universidad.[14]

APQC es un referente que proporciona la taxonomía para la estandarización de los procesos de diversas industrias, entre las que figura la educación superior. La estandarización de los procesos facilita la evaluación comparativa (*benchmarking*) para un análisis y debate objetivo de las actividades.[15]

IWA2:2007 (*International Workshop Agreement*), es un marco de referencia orientado a la gestión de la calidad de la educación superior. Basada en ISO 9001:2000.[16]

PMBOK (*Project Management Body of Knowledge*), orientado a la gestión de proyectos, debe ser usado en los ejercicios de arquitectura empresarial de las universidades para gestionar las iniciativas seleccionadas.[17]

PRINCE 2 (*Projects in Controlled Environments*), apto para la gestión de iniciativas de arquitectura empresarial, provee un conjunto de buenas prácticas para la gestión de proyectos con un manejo de altos niveles de variabilidad e incertidumbre en entornos controlados.[18]

CMM (*Capability Maturity Model*), es un marco de referencia encargado de la evaluación de los procesos. Es usado para definir el grado de madurez de los componentes de los dominios de arquitectura empresarial en una escala de 0 (inexistente) a 5 (implementado y maduro) y la consecuente identificación de brechas. [2]

La Gestión Estratégica de Proveedores, permite la administración y evaluación del rendimiento de los proveedores, a través de la aplicación de mejores prácticas, control de gasto, reducción de riesgos y aprovecha de la generación de valor para el negocio.

El informe anual de Tendencias Mundiales de Capital Humano de *Deloitte*, es una guía que las universidades pueden adoptar para la definición de una estructura organizacional de vanguardia, cuyas estrategias de gestión de capital humano se alineen con la disrupción digital.[19]

A estos referentes se suman otros propios de la gestión de TI tales como:

PETI para la Planeación estratégica de TI que cubra las necesidades de los niveles; estratégicos, directivos y operativos de la institución y garantice el adecuado uso de los recursos, las inversiones tecnológicas y el proceso de implementación de las soluciones tecnológicas. El PETI permite que los responsables de TI cuenten con estrategias de TI embebidas en las estrategias de negocio y realizar la toma de decisiones respecto a las inversiones tecnológicas, la capitalización las innovaciones, la implementación de soluciones de economía de escala, etc. [2]

El marco de referencia COBIT (*Control Objectives for Information and related Technology*) apoya a los líderes de TI de las instituciones educativas con las mejores prácticas de

control y supervisión apuntando a satisfacer las necesidades de los grupos de interés y la generación de valor.[20]

La Biblioteca de Infraestructura de Tecnologías de Información ITIL (*Information Technology Infrastructure Library*), proporciona las mejores prácticas en la gestión de servicios de TI y define acuerdos de niveles de servicio SLA con proveedores de servicios tecnológicos. Los responsables de TI de las instituciones de educación superior pueden adoptar esta guía para: Gestionar Incidencias, Problemas, Cambios, Configuraciones y Versiones de las soluciones tecnológicas y los servicios que nacen de las iniciativas de arquitectura empresarial. [2]

El uso de la arquitectura SOA (*Service Oriented Architecture*) en las instituciones de educación superior, apoyan la implementación de un esquema de interacción e integración con los sistemas de información de externas[21] como CEAACES, SENESCYT, etc. Este modelo propicia la accesibilidad de las aplicaciones web y contenido públicos propio de estas instituciones.

Las universidades al igual que otras organizaciones deben tener una gestión eficiente de sus activos de información. DAMA-DMBOOK es un marco de referencia que proporciona las directrices para la gestión eficiente de la información.[22]

Finalmente, para la protección de datos y seguridades de la institución, debe aplicarse el estándar ISO 27002-2013 para implementar el Sistema de Gestión de las Seguridades de Información. Este estándar proporciona las normas para la selección, implementación y administración de controles de la seguridad de la información y las mejores prácticas de la gestión de la seguridad de la información.[23]

IV. CONCLUSIONES

La solución de problemas empresariales a través de los procesos de Arquitectura Empresarial en el Ecuador es relativamente nueva si lo comparamos con países de la región como Colombia y Costa Rica; por lo tanto, el aporte de este trabajo radica en promover en las universidades del Ecuador implementaciones de arquitectura empresarial.

La adopción de soluciones pertinentes a través de arquitectura empresarial en las universidades ecuatorianas, radica fundamentalmente en el exigente y complejo entorno actual debido a los altos niveles de competitividad, la necesidad de sostenibilidad, las exigencias de los organismos rectores y las tendencias mundiales de la digitalización.

La descripción del proceso de arquitectura empresarial está fundamentada en el marco de referencia TOGAF propuesto por el Open Group. Su elección se basa en el estudio comparativo realizado por Robert Sessions en el que TOGAF obtiene una importante puntuación en la *Integridad del proceso, neutralidad del proveedor y la disponibilidad de información*. Su ventaja frente a FEA, Zachman y Gartner rase debe a los costos de implementación. [2]

Los marcos de referencia complementarios de este trabajo son propuestos a modo de sugerencia. El amplio abanico de posibilidades de buenas prácticas, normativas y modelos

del mercado dejan en libertad de elección a los responsables de la arquitectura empresarial. Sin embargo, en este trabajo se ha procurado seleccionar los referentes que mejor se orientan a la vertical de la educación superior.

Los marcos de referencia complementarios propuestos en el presente trabajo, se correlacionan con las actividades de arquitectura empresarial en los dominios; del negocio, aplicaciones, datos y tecnología y se deben proponer en el mapa de ruta detallado de la fase E de TOGAF

REFERENCIAS

- [1] J. Monsalve, "Referencia de Arquitectura Empresarial," 2015.
- [2] B. Avila, *Formulación de Visionamiento Estratégico de Arquitectura Empresarial aplicado a Entidades de Educación Superior: Un Caso de Aplicación*. PhD thesis, 2016.
- [3] M. Darío Arango Serna, J. Enrique Londoño Salazar, and J. Andrés Zapata Cortés, "ARQUITECTURA EMPRESARIAL – UNA VISIÓN GENERAL," *Revista Ingenierías Universidad de Medellín*, vol. 9174, no. 16, 2010.
- [4] "Business Strategy - Enterprise Architecture."
- [5] Wikidot, "Application Architecture - Enterprise Architecture."
- [6] Wikidot, "Technology Architecture - Enterprise Architecture."
- [7] C. O'Rourke, N. Fishman, and W. Selkow, *Enterprise architecture using the Zachman Framework*. Course Technology, a division of Thomson Learning, 2003.
- [8] J. Schekkerman, *How to survive in the jungle of enterprise architecture frameworks : creating or choosing an enterprise architecture framework*. Trafford, 2004.
- [9] "Federal Enterprise Architecture Framework," 2013.
- [10] R. S. Bittler and G. Kreizman, "Gartner Enterprise Architecture Process: Evolution 2005," *Gartner Enterprise Architecture Process*, no. October, p. 12, 2005.
- [11] The Open Group, *TOGAF Versión 9.1*. 2011.
- [12] J. Laviña Orueta dir., L. Mengual Pavón dir., and F. Telefónica, *Libro blanco de la universidad digital 2010*. 2008.
- [13] M. E. Porter and M. E. Rosas-Sánchez, *Estrategia competitiva : técnicas para el análisis de los sectores industriales y de la competencia*. Continental, 2000.
- [14] T. d. Marmol, *BCG growth-share matrix*.
- [15] APQC, "Education Process Classification Framework 7.0.5," 2016.
- [16] IWA, "Guía de aplicación -IWA 2 - Sistemas de gestión de la calidad para organizaciones educativas," pp. 1–100, 2002.
- [17] Project Management Institute, *PMBOK: Guía de Fundamentos de Dirección de Proyectos*. 2013.
- [18] O. o. G. Great Britain. Office of Government Commerce. and O. o. G. Commerce, *Managing successful projects with PRINCE2*. TSO, 5th ed., 2009.
- [19] Deloitte, "Tendencias Mundiales de Capital Humano 2017."
- [20] ISACA, *COBIT 5 - Un Marco de Negocio para el Gobierno y la Gestión de TI de la Empresa*. 2012.
- [21] I2B, "¿Qué se entiende por SOA, y cuáles son sus beneficios? | I2B Intelligence to Business."
- [22] M. Mosley, "DAMA DMBOK Functional Framework," *Dama-Dmbok*, vol. 3.02, pp. 1–19, 2008.
- [23] ISO, "ISO 27002 en español," *Normativa ISO*, vol. 2013, 2013.

Recibido: 3 de mayo de 2018

Aceptado: 15 de junio de 2018



DetECCIÓN DE FILTRACIONES DE PRIVACIDAD: UNA VISIÓN DEL ESTADO DEL ARTE

DETECTION OF PRIVACY LEAKS IN SOCIAL NETWORKS: A VISION OF THE STATE OF ART

Diana Marcela Romero Córdova^{1*}

Unidad Académica de Tecnologías de la Información y la Comunicación,
Carrera de Ingeniería de Sistemas, Universidad Católica de Cuenca¹

*dmromeroc@ucacue.edu.ec

DOI: https://doi.org/10.26871/killkana_tecnica.v2i1.289

Resumen

Cuando se habla de redes sociales, se expresan todas las ventajas que estas brindan al usuario en cuanto a la comunicación e interacción con el mundo, además se reconoce que son una fuente de extracción y uso de datos con diversos fines, aunque algunos con objetivos ilícitos. Las redes sociales se han popularizado y las personas les dedican diariamente tiempo para revisar las principales notificaciones recibidas, así como comunicarse, publicar, realizar comentarios o colocar etiquetas. La necesidad de controlar la privacidad es fundamental, pero no son suficientes las configuraciones que se puedan realizar para evitarlo. El usuario coloca en sus publicaciones mensajes que contienen elementos propios del lenguaje natural, pero que resultan difíciles de interpretar por un software. Para realizar una búsqueda de elementos considerados privados que se encuentren en los textos publicados o que puedan inferirse por la revisión total de la información visible publicada en el perfil del usuario, es necesario procesar el texto para limpiarlo, normalizarlo y clasificarlo. La información privada puede ser fácilmente inferida, con solo asociar las diferentes publicaciones textuales que se compartieron por el usuario u otra persona. Resulta oportuno hacer uso de las técnicas de la minería de texto para evitar la publicación de información privada en el perfil que se comparte con el grupo de amigos de la red social.

Palabras clave: filtraciones, minería de texto, privacidad, redes sociales.

Abstract

When talking about social networks, all the advantages they offer to the user in terms of communication and interaction with the world are expressed, and it is also recognized that they are a source of extraction and use of data for various purposes, although some with illicit purposes. Social networks have become popular and people dedicate daily time to review the main notifications received, as well as communicate, publish, comment or place labels. The need to control privacy is fundamental, but the configurations that can be made to avoid it are not enough. The user posts messages that contain elements of natural language, but that are difficult to interpret by software. To perform a search for items considered private that are found in the published texts or that can be inferred by the total revision of the visible information published in the user's profile, it is necessary to process the text to clean it, normalize it and classify it. Private information can easily be inferred, just by associating the different textual publications that were shared by the user or another person. It is opportune to make use of text mining techniques to avoid the publication of private information in the profile that is shared with the group of friends of the social network.

Key words: leaks, privacy, social networks, text mining.

I. INTRODUCCIÓN

La evolución de la tecnología en los últimos años ha permitido que personas de diferentes edades, en distintas ubicaciones geográficas y utilizando un dispositivo conectado a internet pueda comunicarse con el mundo entero de una manera dinámica, permitiendo el intercambio de información y relacionando personas a través de las redes sociales. [1] [2]

Al estar conectados, los usuarios han podido identificar no solo las múltiples posibilidades y ventajas del uso de las redes sociales, también las innumerables formas en las que se expone su información personal, y los riesgos que conllevan. Actualmente se regulan los contenidos publicados en redes sociales o medios electrónicos [1] en varios países. El Ecuador incluye en su código orgánico integral penal COIP artículos que protegen los datos e información

de los ciudadanos, incluyendo los que se publican en las redes sociales, en búsqueda de una mejor convivencia entre sus habitantes.

Las redes sociales se han convertido en el objetivo de muchos investigadores. En este documento el objetivo consiste en revisar el estado actual de las principales formas de control de la información textual que se publica en las redes sociales, que con el uso de diferentes técnicas permitan identificar elementos de privacidad. Se han convertido en una fuente de análisis los contenidos y comportamientos de los usuarios. [3]

En la sección II se define la privacidad desde diferentes puntos de vista y su importancia, siendo necesario el control de los diferentes tipos de datos que se publican. En la sección III se indican las recomendaciones a los usuarios de las redes sociales para evitar perder datos privados. En la sección IV se describen las técnicas de protección de datos en las redes sociales. La sección V tiene el detalle de las principales técnicas de extracción de información de las redes sociales para su análisis, dónde la técnica de la minería de textos demuestra su importancia al momento de procesar los textos en lenguaje natural, a continuación en la sección VI se explica el proceso de clasificación automática de los textos realizada con minería de textos y se finaliza con la sección VII donde se presentan las conclusiones del documento.

II. LA PRIVACIDAD

Según la Real Academia de la lengua española se define la privacidad como "ámbito de la vida privada que se tiene derecho a proteger de cualquier intromisión".

Mark Zuckerberg dijo que "la privacidad es un bien del pasado y que, tecnológicamente hablando, el mundo actual es un lugar en el que no tiene cabida. Creo que son visiones profundamente erradas, pues la privacidad es un asunto fundamental para la libertad de las personas, y también estoy convencido de que se pueden construir herramientas tecnológicas con un mejor tratamiento para la privacidad de los usuarios" [4]

En [5] se define a la privacidad como el derecho de un individuo al secreto de su vida. Estudios realizados por [6], [7], [8] mencionan la importancia de proteger los datos de las personas y empresas. La información y los datos aportan gran valor a las empresas, tanto que son considerados parte de sus activos. [8]

Al hacer uso de las redes sociales un usuario coloca en ellas gran cantidad de datos, que pueden clasificarse en dos grupos:

- Datos personales: Los que el usuario coloca en el perfil que se crea, estos datos quedan expuestos dependiendo de las preferencias de privacidad configuradas por el usuario. Además se exponen las características físicas, como el color del cabello, forma del rostro; datos académicos y de contacto. [4] Las publicaciones que

incluyen fotografías o videos facilitan esta identificación.

- Datos sensibles: Aquellos datos personales que revelen el origen racial del usuario, sus ideologías, creencias religiosas, afiliación política, preferencias sexuales, etc. [4] en muchos casos se incluyen los datos empresariales o de negocios como cuentas bancarias, proveedores, etc.

La difusión de comentarios en los perfiles de las redes sociales de los usuarios son visibles para todos sus contactos [9]. Estos comentarios pueden ser leídos y asociados por una persona que pertenezca a la red y que esté interesada en descubrir información privada. Con un proceso mental de análisis de los textos publicados, podrá inferir información identificando elementos que ponen en riesgo la privacidad del usuario. [10]

III. PRINCIPALES RECOMENDACIONES A LOS USUARIOS DE REDES SOCIALES

Las principales recomendaciones que se han emitido a los usuarios de las redes sociales inician con la modificación voluntaria de las configuraciones de privacidad de su perfil. Sin embargo, no es suficiente. La privacidad se pierde el momento de la creación de la cuenta, porque datos como correo electrónico, número de teléfono celular, fecha de nacimiento, nombres y apellidos son obligatorios para su autenticación. La información del usuario pasa a formar parte de las grandes bases de datos del proveedor de la red social y no pueden ser eliminados. [11]

El derecho al olvido es una de las regulaciones que se están proponiendo en Europa, con el objetivo de poder eliminar definitivamente de los grandes servidores los datos que los usuarios han compartido. [8]

Algunos autores mencionan la necesidad de integrar en el currículo de formación de los jóvenes el fomento del buen uso de las redes sociales. [12] Concienciando así al uso responsable de las redes sociales y la información que en ellas se comparte. Los mensajes de odio que se encuentran en las redes sociales pueden influir principalmente en los jóvenes, quienes son más propensos a caer en provocaciones. El desarrollo de habilidades críticas es un requerimiento para el uso ético de estas redes. [13]

Las redes sociales son usadas en diferentes áreas profesionales, por ejemplo en la médica, con claras ventajas en la creación de perfiles, intercambio de artículos científicos y comunicación con los pacientes. Sin embargo, [14] menciona la importancia de mantener un perfil profesional modesto, con la creación de contenido responsable y honesto.

Otra recomendación importante tiene que ver con el crecimiento de la red. Los miembros que la integran deben formar parte del grupo de amigos, familiares o conocidos reales. La aceptación o la integración de nuevos miembros para pertenecer al grupo de amigos debe realizarse con precaución. Muchos de los perfiles son falsos y un gran porcentaje tienen como objetivo formar parte de una red

social con intenciones de acoso, estafa o robo de identidad. [15]

IV. EXTRACCIÓN DE INFORMACIÓN TEXTUAL DE LAS REDES SOCIALES

Las redes sociales generan diariamente 2,5 quintillones de bytes en datos [16] de tipo no estructurado¹. Estas publicaciones se realizan en lenguaje natural² fácilmente entendido por las personas pero difícil de interpretar por un software, razón por lo que el uso de técnicas especializadas para el análisis, procesamiento y organización de este tipo de datos es necesario.

La necesidad de identificar elementos que pueden revelar información considerada privada en un perfil de redes sociales antes de su publicación, dan paso al estudio de las principales técnicas para identificar los textos, procesarlos, categorizarlos e identificar relaciones entre ellos. La ciencia que permite el proceso mencionado es la minería de textos [17], que facilita la extracción de información a partir de textos publicados [16], descubriéndose conocimiento a través de la identificación de patrones ocultos en los mismos. [18]

La minería de textos (MT) puede ser confundida con la minería de datos (MD)³, pero a diferencia de la segunda, la MT extrae información clara y explícita de los textos publicados no estructurados [19] [17] [16].

En la MT se abordan problemas como la categorización (clasificación supervisada)⁴ y el agrupamiento (clasificación no supervisada)⁵ entre otros.

El procesamiento de lenguaje natural (PLN) es una rama de la Inteligencia Artificial cuyo objetivo es conseguir que un software comprenda el texto escrito por una persona, ayudando a la interacción entre el humano y la computadora. [20] [10]. Siri de IOS, Cortana de Microsoft, el traductor de google entre otros varios, son productos incorporan algoritmos de procesamiento de lenguaje natural. [20]

El análisis de sentimiento de los mensajes publicados en las redes sociales, se pueden procesar combinando las técnicas de minería de texto y procesamiento de lenguaje natural. [21] [22] Las opiniones que se expresen en las redes sociales aportan en la identificación de la reputación online, [23] [24] que es aprovechada por instituciones gubernamentales, empresariales, educativas, financieras, etc.

¹Los datos no estructurados, generalmente son datos binarios que no tienen estructura interna identificable. Es un conglomerado masivo y desorganizado de varios objetos que no tienen valor hasta que se identifican y almacenan de manera organizada.

²El lenguaje natural es el medio que se utiliza para la comunicación cotidiana con las demás personas por ejemplo el español o el inglés.

³Es un campo de la estadística y ciencias de la computación que intenta descubrir patrones en grandes volúmenes de datos estructurados

⁴Son aquellos que a partir de un conjunto de datos de entrenamiento, se intenta asignar una clasificación a un segundo conjunto de datos. Se preocupa de asignar etiquetas que indican a qué categoría o categorías corresponde el documento

⁵Sin poseer un conjunto de datos de entrenamiento se parte de las propiedades de los ejemplos y se intenta agruparlos de acuerdo a su similitud. Consiste en la división de los datos en grupos de objetos similares

El reconocimiento de entidades nombradas⁶ es una de las técnicas de minería de textos que contribuye a mejorar la tarea en la minería de opinión. [25] El enfoque de aprendizaje supervisado permite la búsqueda de respuestas y realiza la categorización de textos de manera automática.

V. PROCESO DE CLASIFICACIÓN AUTOMÁTICA DE LOS TEXTOS

En esta sección se presentan aspectos importantes de la clasificación automática de los textos, que sirven para identificar y clasificar cada uno de los mensajes que se encuentran publicados en el perfil de la red social que pudo haber sido compartida por el usuario propietario o un amigo que pertenece a su red, realizando una búsqueda de filtraciones de elementos privados.

1. Pre procesamiento de los textos

Las operaciones de pre procesamiento de textos se centran en la identificación, extracción, refinamiento y adición de características a los textos publicados en las redes sociales. De tal forma que se descarten todos aquellos elementos no representativos. [26]

Una primera etapa consiste en separar el texto en palabras, frases, símbolos u otra unidad sintáctica llamada token [25]. Estos tokens identificados serán la entrada para las etapas posteriores del análisis. En las redes sociales cada mensaje deberá ser dividido en tokens.

La conversión de mayúsculas a minúsculas elimina uno de los atributos que permite identificar a las entidades nombradas, la letra mayúscula [27] esta etapa se denomina normalización⁷. Para continuar es necesaria la eliminación de preposiciones, artículos y conjunciones.

La eliminación del ruido es la tercera etapa, consiste en la eliminación de los espacios innecesarios y signos de puntuación de los textos. Luego se convertirá el texto en un documento de formato plano, es decir carecerá de estilo (negrita, subrayado, un único tipo de letra). Esta etapa es importante porque los mensajes escritos por los usuarios por lo general poseen urls, emoticones, hashtags, etc., propios del lenguaje natural. El objetivo es precisamente eliminar estos elementos que distorsionan el texto y facilitar su procesamiento. [22].

Para concluir con este proceso se eliminan los stop words, que son palabras con poco aporte semántico.

2. Identificación de nombres propios

En esta etapa se necesita de técnicas de análisis sintáctico⁸ que identifican los verbos que sirven de unión entre nombre propios (nombres de instituciones, de personas organizaciones, etc.). Es fundamental que se identifiquen estos nombres en los textos publicados para poder encontrar relaciones entre ellos.[16]

⁶Nombres de personas, instituciones, lugares, productos

⁷Las palabras se agrupan según su familia léxica

⁸Relaciones de concordancia y jerarquía que guardan las palabras cuando se agrupan en oraciones

3. Representación de los documentos mediante un modelo

Generalmente en esta etapa se utiliza un modelo vectorial para realizar la representación de los documentos, donde cada texto o documento representa un vector[16].

4. Categorización automática

En primer lugar a cada documento es necesario asignarle una categoría, o a un mismo documento asignarle más de una categoría. De ser necesario se clasifica cada documento en una categoría mediante verdadero o falso. Por último se indica la probabilidad que tiene cada documento de pertenecer a una categoría[28].

5. Relaciones entre términos y conceptos

En esta última etapa se realizan la extracción de términos y conceptos y las relaciones existentes entre ellos[28].

Al terminar de ejecutar todo el proceso descrito, se obtiene un corpus de texto estructurado,

VI. COMPONENTES DE SOFTWARE DE CLASIFICACIÓN

Las técnicas de de clasificación de la minería de textos consisten en asignar a las categorías previamente definidas diferentes objetos.[26] Los clasificadores pueden construirse a partir de aprendizaje de máquina. Cuando esto sucede, un proceso inductivo examina las características de los documentos previamente clasificados e infiere las condiciones que los nuevos textos deben cumplir para ser clasificados en una u otra categoría. Por ejemplo los contenidos de las publicaciones en las redes sociales serán clasificados como datos privados o comunes. Al realizar la comparación de las categorías identificadas en el corpus de entrenamiento con las categorías seleccionadas automáticamente es posible identificar la calidad del clasificador. [26]

A continuación se describen los componentes de un clasificador que usa aprendizaje de máquina.

1. Categorías

Para iniciar con el entrenamiento, previamente se deben identificar y registrar las categorías en las que se deben clasificar los diferentes componentes del corpus. Estas categorías son revisadas por un experto. Una manera de comprobar la efectividad del clasificador, será comparar los textos clasificados con los datos del entrenamiento.[26]

2. Conjunto de entrenamiento

Los datos utilizados para construir los clasificadores son tomados de las redes sociales a través de sus api. Es necesario identificar las categorías en las que debe basarse el clasificador para realizar su trabajo.

3. Selección de características y algoritmo de clasificación

En esta sección es importante entender las necesidades del negocio, es decir analizar el contenido textual de las publicaciones en las redes sociales en la primera fase denominada conocer el negocio.

A continuación la fase de comprensión de los datos se deben considerar tres aspectos: errores ortográficos, incompatibilidad de caracteres y archivos duplicados.

La tercera fase es la de preparación de los datos, que consiste el procedimiento descrito anteriormente denominado pre procesamiento de los textos.

La cuarta fase se refiere al modelado, en la que se utilizan distintos algoritmos de clasificación, pudiendo se utilizados: Naïve Bayes, al ser un clasificador probabilístico su desempeño depende de la cantidad de instancias empleadas para estimar las probabilidades. El algoritmo KNN (k-ésimo vecino más cercano) no construyen una representación declarativa explícita de las categorías sino que computan directamente la similitud entre el documento a ser clasificado y los documentos de entrenamiento. SVM (Máquina de soporte vectorial) binario, es usado cuando se tienen solo dos categorías, por ejemplo positiva y negativa.[26]

Finalmente es importante realizar la evaluación de los datos obtenidos por cada uno de los algoritmos utilizados, a fin de comparar sus resultados y realizar las iteraciones que fueran necesarias hasta encontrar los resultados que indiquen la eficiencia del proceso.

En la literatura se identifican gran cantidad de herramientas o softwares que permiten la minería y clasificación automática de los textos utilizando herramientas de aprendizaje supervisado. En [29] se mencionan los software que se encuentran en desarrollo: Software de Asistente de Creación de Corpus, Sistema de Mapeo de Datos, Sistema de detección de similitudes en archivos de código fuente, que permiten el reconocimiento del contexto de los textos.

VII. CONCLUSIONES

Gran cantidad de investigaciones sobre las redes sociales y sus seguridades están enfocadas en la detección de la privacidad de los usuarios a partir de la estructura de la red. Este trabajo propone utilizar las técnicas de minería de textos para realizar la estructuración de sus contenidos. Para futuras investigaciones es importante diseñar un modelo que permita caracterizar e identificar los elementos considerados privados por parte del usuario previo a una clasificación y relación de las entidades identificadas en el texto. Lo que contribuirá a una propuesta de seguridad para el usuario a partir de datos textuales publicados.

REFERENCIAS

- [1] J. P. Cardoso, "Democracia y redes sociales," *Universidad Verdad*, no. 72, pp. 297–311, 2017.
- [2] E. Amesti, L. Estrada, and D. Rey, "Inteligencia de negocios y redes sociales//business intelligence and

- social networks,” *Marketing Visionario*, vol. 2, no. 2, pp. 110–124, 2014.
- [3] L. F. Hurtado, F. Pla, M. Giménez, and E. S. Arnal, “Elirf-upv en tweetlid: Identificación del idioma en twitter,” in *TweetLID@ SEPLN*, pp. 35–38, 2014.
- [4] S. Flores and E. Estefanía, “Análisis de la privacidad y transparencia del internet,” B.S. thesis, PUCE, 2017.
- [5] P. E. Núñez Fiallos, “La ley de comunicación y la violación del derecho a la privacidad individual en las redes sociales,” B.S. thesis, Universidad Técnica de Ambato, Facultad de Jurisprudencia y Ciencias Sociales, Carrera de Derecho, 2018.
- [6] T. Moreno-Becerra, C. Gajardo-León, and E. Parra-Ortiz, “Privacidad: cómo se entiende y se gestiona en facebook. estudio de caso de jóvenes chilenos,” *Revista Latina de Comunicación Social*, no. 71, pp. 715–729, 2016.
- [7] M. Toscano, “Sobre el concepto de privacidad: la relación entre privacidad e intimidad,” *Isegoría*, no. 57, pp. 533–552, 2017.
- [8] M. Á. Caro, *Derecho al olvido en internet: el nuevo paradigma de la privacidad en la era digital*. Editorial Reus, 2015.
- [9] L. R. García, J. R. M. Benedito, *et al.*, “Perspectiva de los jóvenes sobre seguridad y privacidad en las redes sociales,” *Revista ICONO14 Revista científica de Comunicación y Tecnologías emergentes*, vol. 14, no. 1, pp. 24–49, 2016.
- [10] E. Cambria and B. White, “Jumping nlp curves: A review of natural language processing research,” *IEEE Computational intelligence magazine*, vol. 9, no. 2, pp. 48–57, 2014.
- [11] L. Tello, “Intimidad y «extimidad» en las redes sociales. las demarcaciones éticas de facebook,” *Comunicar*, vol. 21, no. 41, 2013.
- [12] E. Argente, E. Vivancos, J. Alemany, and A. García-Fornes, “Educar en privacidad en el uso de las redes sociales,” *Education in the Knowledge Society*, vol. 18, no. 2, p. 107, 2017.
- [13] I. A. Martínez, “La cultura sobre seguridad informática en las redes sociales: el caso de los estudiantes de la preparatoria de san diego cuentla, méxico/the culture on information security in social networks: the case of students of san diego cuentla high school, mexico,” *RICSH Revista Iberoamericana de las Ciencias Sociales y Humanísticas*, vol. 6, no. 11, pp. 20–30, 2017.
- [14] M. Rodríguez-Socarrás, J. Gómez-Rivas, M. Álvarez-Maestro, L. Tortolero, M. Ribal, M. G. Sanz, and M. Rouprêt, “Adaptación al español de las recomendaciones para el uso apropiado de las redes sociales en urología de la european association of urology (eau),” *Actas Urológicas Españolas*, vol. 40, no. 7, pp. 417–423, 2016.
- [15] E. Zheleva, E. Terzi, and L. Getoor, “Privacy in social networks,” *Synthesis Lectures on Data Mining and Knowledge Discovery*, vol. 3, no. 1, pp. 1–85, 2012.
- [16] A. I. Valero Moreno, “Técnicas estadísticas en minería de textos,” 2017.
- [17] J. Pino-Díaz *et al.*, “Minería de textos,” *Ingeniería del conocimiento biomédico y del producto, I+ D en investigación traslacional. Master Universitario Investigación Traslacional y Medicina Personalizada (Transmed) de la Universidad de Granada.*, 2016.
- [18] A. R. Blanco, A. S. Cuevas, E. G. Martínez, and W. H. Mazo, “Modelo de representación de textos basado en grafo para la minería de texto,” *Ciencias de la Información*, vol. 46, 2015.
- [19] A. A. Consuegra, Y. M. Salazar, J. H. García, and D. H. Vizcaino, “Minería de texto como una herramienta para la búsqueda de artículos científicos para la investigación,” *INVESTIGACIÓN Y DESARROLLO EN TIC*, vol. 7, no. 1, pp. 14–20, 2017.
- [20] Á. García Gutiérrez *et al.*, “Machine learning en bases de datos de lenguaje natural,” B.S. thesis, 2016.
- [21] E. Martínez Cámara, M. T. Martín Valdivia, J. M. Perea Ortega, and L. A. Ureña López, “Técnicas de clasificación de opiniones aplicadas a un corpus en español,” *Procesamiento del Lenguaje Natural*, no. 47, 2011.
- [22] S. Estévez-Velarde and Y. A. Cruz, “Evaluación de algoritmos de clasificación supervisada para el minado de opinion en twitter,” *Investigación Operacional*, vol. 36, no. 3, pp. 194–205, 2015.
- [23] M. Del-Fresno-García, “Cómo investigar la reputación online en los medios sociales de la web 2.0,” *Cuadernos de comunicación Evoca*, vol. 5, no. 1, pp. 29–33, 2011.
- [24] J. Fernández, Y. Gutiérrez, J. M. Gómez, and P. Martínez-Barco, “Social rankings: análisis visual de sentimientos en redes sociales,” *Procesamiento del Lenguaje Natural*, vol. 55, pp. 199–202, 2015.
- [25] L. G. Fernández, “Metodología para el reconocimiento de entidades nombradas en mensajes cortos,” 2017.
- [26] P. Santana Mansilla, R. Costaguta, and D. Missio, “Aplicación de algoritmos de clasificación de minería de textos para el reconocimiento de habilidades de tutores colaborativos,” *Inteligencia Artificial. Revista Iberoamericana de Inteligencia Artificial*, vol. 17, no. 53, 2014.
- [27] A. Ritter, S. Clark, O. Etzioni, *et al.*, “Named entity recognition in tweets: an experimental study,” in *Proceedings of the conference on empirical methods in natural language processing*, pp. 1524–1534, Association for Computational Linguistics, 2011.
- [28] J. A. Valero-Medina, C. D. Dallos-Bustos, and I. Lizarazo, “Un nuevo enfoque para la clasificación de imágenes multispectrales basado en complejos cartesianos,” *DYNA*, vol. 85, no. 204, pp. 28–37, 2018.
- [29] J. J. Castillo, M. E. Cardenas, A. Curti, O. Casco, M. Navarro, N. A. Hernández, and M. Velazco, “Desarrollo de sistemas de análisis de texto,”

in XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017, ITBA, Buenos Aires), 2017.

Recibido: 3 de mayo de 2018

Aceptado: 15 de junio de 2018





Instructivo para la preparación de manuscritos en la revista Killkana Técnica

Instructions for the preparation of manuscripts in the Killkana Técnica Journal

Nombre1 Apellido1^{*1} y Nombre2 Apellido2²

¹ Facultad, Programa o Grupo de Investigación, Institución

² Facultad, Programa o Grupo de Investigación, Institución

*nombre_autor1@correo.com

Resumen

La finalidad de éste documento es proporcionar las instrucciones y normas para la presentación de artículos en la Revista Killkana Técnica. Es ejemplo del diseño editorial deseado (incluido este resumen) y puede usarse como referencia en general para la buena escritura de artículos de investigación. Se dan reglas sobre citas, ecuaciones, unidades, figuras, tablas, abreviaturas y acrónimos. También se orienta la redacción de las partes de un artículo: resumen introducción, cuerpo del documento, referencias y biografías de los autores, etc. El resumen debe tener como mínimo 200 palabras y como máximo 250 palabras y no puede contener ecuaciones, figuras, siglas, tablas ni referencias. Debe presentar en forma resumida la problemática, describir el objetivo general, la metodología desarrollada, los principales resultados, impactos y las conclusiones más relevantes del trabajo.

Palabras clave: El autor debe proporcionar palabras clave (en orden alfabético), un mínimo de 3 y un máximo de 6, que ayuden a identificar los temas o aspectos principales del artículo.

Abstract

Debe contener la traducción del resumen en idioma Inglés (traducción con alta calidad).

Key words: Contiene la traducción de las palabras claves al idioma Inglés.

I. INTRODUCCIÓN

LA revista KILLKANA TÉCNICA de la Universidad Católica de Cuenca, tiene como propósito principal estimular los esfuerzos científicos, tecnológicos, docentes y de extensión, de los Miembros de la Comunidad Científica Nacional e Internacional en el área de las ingenierías. Pretende poner a disposición de docentes, investigadores y profesionales un medio de promoción y difusión que brinde la oportunidad de dar a conocer el fruto de sus trabajos y les permita expresar sus opiniones respecto a cualquier actividad fundamental en sus áreas de experticia.

Este documento proporciona un ejemplo de diseño de edición de un artículo para la revista KILLKANA TÉCNICA. Es una guía para la redacción que, conjuntamente con las instrucciones para el uso de la clase *Cómo usar la clase de documento L^AT_EX 'killkanaT.cls' de la revista Killkana Técnica*, proveen de la información necesaria para la elaboración de un artículo. Contiene indicaciones del formato de edición y de los estilos empleados en la revista. Se dan reglas de estilo sobre redacción, secciones, citas, ecuacio-

nes, figuras, tablas, abreviaturas, referencias bibliográficas, acrónimos, etc.

Para producir un artículo para la revista KILLKANA TÉCNICA se requiere del archivo de la clase *killkanaT.cls* y el de estilo *killkanaT.sty*, los cuales se distribuyen a través de la licencia LPPL (L^AT_EX Project Public License) y pueden obtenerse a través del contacto con el editor de la revista mediante el correo electrónico: killkana.editorial@ucacue.edu.ec o en la página web de la revista <http://www.killkana.ucacue.edu.ec>.

II. CRITERIOS GENERALES DE CLASIFICACIÓN DE LOS ARTÍCULOS

Diferentes tipos de artículos son aceptados para ser sometidos a la evaluación de pares especialistas en el área determinada del trabajo.

A. Artículo científico y tecnológico

Documento que presenta, de manera detallada y sistemática, los resultados originales de proyectos terminados de investigación.

B. Artículo de revisión

Documento resultado de una investigación terminada donde se analizan, sistematizan e integran los resultados de investigaciones publicadas o no publicadas, sobre un campo de la ciencia o la tecnología, con el fin de dar cuenta de los avances y las tendencias de desarrollo. Se caracterizan por presentar una cuidadosa revisión bibliográfica de por lo menos 50 referencias.

C. Artículo corto

Documento breve que presenta resultados originales preliminares o parciales de una investigación científica o tecnológica, que por lo general requieren de una pronta difusión.

D. Reporte de casos

Documento que presenta los resultados de un estudio sobre una situación particular con el fin de dar a conocer las experiencias técnicas y metodológicas consideradas en un caso específico. Incluye una revisión sistemática comentada de la literatura sobre casos análogos.

E. Artículo de reflexión

Documento que presenta resultados de investigación terminada desde una perspectiva analítica, interpretativa o crítica del autor, sobre un tema específico, recurriendo a fuentes originales. Este tipo de documentos se conocen en el ámbito académico como ensayos y en general son escritos que mediante una estructura identificable exponen con claridad un tema, desarrollan unas implicaciones y presentan unas conclusiones.

III. CONTENIDOS DE UN ARTÍCULO

La estructura generalmente utilizada en la conformación del cuerpo de un artículo consta de seis apartes importantes: introducción, materiales y/o métodos, resultados, discusión, conclusiones, agradecimientos y fuentes o referencias bibliográficas.

Se aconseja cuidar la ortografía, la sintaxis y la semántica empleando un corrector automático de texto. Las oraciones deben estar completas, con sentido claro e inequívoco y continuidad entre ellas, así como entre párrafos.

El cuerpo del artículo debe contener las siguientes características [1]:

- 1) Resumen - palabras clave -
- 2) Cuerpo del documento:
 - a) Introducción.
 - b) Marco teórico.
 - c) Metodología y/o materiales. Materiales sólo en el caso de un artículo experimental y/o de laboratorio.
 - d) Resultados.
- 3) Conclusiones y recomendaciones.
- 4) Referencias.

A. Título y nombres de los autores

Comenzaremos, aunque parezca obvio y no se le de mayor importancia, con el título del trabajo. Un investigador interesado en el área de trabajo del autor primeramente encontrará artículos a través de las palabras claves, filtrando de entre muchos títulos que emerjan de la búsqueda. Un posible lector se decidirá por su artículo si el título le parece llamativo e interesante, para luego pasar a leer el resumen. Aquí es donde estriba la importancia de pensar cuidadosamente las palabras que conforman el título. El profesor de la Universidad de Texas en Austin, Harry L. Swinney [2] recomienda realizar una lista de palabras que describan el trabajo, y emplear ésta lista para formar un título corto, de no más de 10 palabras. Evite jergas, palabras con doble significado, y palabras huecas o frases tales como: nueva, novedosa, estudio de, investigación de, exploración de, precisión, alta resolución, eficiente, poderoso.

Luego del título se indica el nombre de los autores, los cuales deben tener un formato similar al de las citas; nombre y apellido del autor, o inicial de los nombres y apellidos, dependiendo de la cantidad de autores y el espacio disponible. Los autores deben estar separados por comas y para el último colocar "y". Omita el título o profesión de los autores.

B. Resumen y palabras claves

El resumen representa una síntesis breve de los contenidos del estudio que permite que los lectores conozcan las generalidades y resultados de la investigación. Como se indicó anteriormente debe tener como mínimo 200 palabras y como máximo 250 palabras y no puede contener ecuaciones, figuras, siglas, tablas ni referencias.

Un resumen cuidadosamente escrito es crucial, porque la gran mayoría de potenciales lectores saltará al próximo artículo después de leer unas cuantas líneas de nuestro resumen. Solamente una pequeña fracción de los lectores del resumen leerán el artículo completo, sea impreso, online o descargado en pdf [2].

El título y el resumen, pese a que aparecen de primero en un artículo, es lo último que se escribe.

Las palabras claves son términos que identifican al tipo de investigación o trabajo realizado y son útiles para ayudar a los indexadores y motores de búsqueda a encontrar los reportes o documentos pertinentes. Se debe escribir un mínimo de tres (3) y un máximo de (6) palabras clave (en orden alfabético).

C. Cuerpo del documento

Cada parte del artículo se dividirá y subdividirá en el grado necesario aprovechando los órdenes de división y encabezamientos que posee \LaTeX : secciones (section), sub-secciones (subsection) y sub-sub-secciones (subsubsection). También se puede hacer uso de los entornos de listas como `enumerate` (lista enumerada) e `itemize` (lista por ítems). Se recomienda revisar el documento *Como usar la clase de documento*

\LaTeX 'killkanaT.cls' de la revista Killkana Técnica, sección VI.

\LaTeX enumera automáticamente las citas, figuras y tablas, sin embargo al incluir las referencias relevantes para su trabajo debe verificar que ellas estén referidas en algún lugar del texto.

1. Introducción

La introducción abarca los antecedentes (tratados de manera breve, concreta y específica), el planteamiento del problema (objetivos y preguntas de investigación, así como la justificación del estudio), un sumario de la revisión de la literatura, el contexto de la investigación (cómo, cuándo y dónde se realizó), las variables y los términos de la investigación, lo mismo que las limitaciones de ésta. Es importante que se comente la utilidad del estudio para el campo académico y profesional.

2. Marco teórico o antecedentes

En este apartado se incluyen y comentan las teorías que se manejaron y los estudios previos que se relacionan con la investigación actual. Se recomienda ser breve y conciso. Algunos autores incluyen esta parte en la introducción.

3. Metodología

En esta parte del artículo se describe cómo fue llevado a cabo el trabajo, e incluye, dependiendo del tipo de investigación realizada, el enfoque; el contexto; casos, universo y muestras; diseño utilizado; procedimiento; y proceso de recolección de datos.

4. Resultados

Los resultados son producto del análisis de los datos. Compendian el tratamiento estadístico y analítico que se dio a los datos. Regularmente el orden es [1]: a) análisis descriptivos de los datos, b) análisis inferenciales para responder a las preguntas o probar hipótesis (en el mismo orden en que fueron formuladas las hipótesis o las variables). Se recomienda que primero se describa de manera breve la idea principal que resume los resultados o descubrimientos, y posteriormente se reporten con detalle los resultados. Es importante destacar que en este apartado no se incluyen conclusiones ni sugerencias.

D. Conclusiones y recomendaciones

Discusiones, recomendaciones, sugerencias, limitaciones e implicaciones. En esta parte se debe [1]:

- Derivar las conclusiones, las mismas que deben seguir el orden de los objetivos de la investigación.
- Señalar las recomendaciones para otros estudios o derivaciones futuras relacionadas con su trabajo.
- Generalizar los resultados a la población.
- Evaluar las implicaciones del estudio.
- Relacionar y contrastar los resultados con estudios existentes.

- Reconocer las limitaciones de la investigación (en el diseño, muestra, funcionamiento del instrumento, alguna deficiencia, etc., con un alto sentido de honestidad y responsabilidad).
- Destacar la importancia y significado de todo el estudio.
- Explicar los resultados inesperados.

En la elaboración de las conclusiones se debe evitar repetir lo dicho en el resumen.

E. Referencias

Son las fuentes primarias utilizadas por el investigador para elaborar el marco teórico u otros propósitos; se incluyen al final del reporte, deben estar enumeradas según el orden de aparición en el documento y siguiendo las normas del *Institute of Electrical and Electronics Engineers* (IEEE).

En la Internet se encuentra suficiente información sobre el estilo de la IEEE, documentos en pdf como [Estilo de Referencias IEEE](#) de la Facultad de Derecho de la UNAM o el [FORMATO IEEE: Estilo y referencias bibliográficas](#) de la Universidad de Málaga, pueden consultarse.

1. Artículos

Nombre o inicial del nombre los autores seguido del apellido, título del artículo entre comillas dobles, nombre de la revista en letras itálicas, volumen (vol.), número (no.), las páginas del artículo (pp) y finalmente mes (abreviado) y año. Ejemplos:

[1] I. Waller & R. Kapral, "Spatial and temporal structure in systems of coupled nonlinear oscillators", *Phys. Rev A*, vol. 30, no. 4, pp. 2047–2055, Oct. 1984.

[2] J. M. S. Nistal, "El tesoro CINDOC de topónimos", *REDIAL: revista europea de información y documentación sobre América Latina*, no. 1, pp. 105-126, 1992.

2. Libros

Nombre o inicial del nombre de los autores seguido del apellido, título del libro (en letras itálicas), seguido del número de la edición (abreviado). País y nombre de la editorial, mes y año de la publicación, capítulo y páginas (abreviadas). Ejemplos:

[3] M. A. Sanjuán & G.D.D.N. Lineal, *Caos, complejidad e interdisciplinariedad. Pensar como un economista*, (Homenaje al profesor Andrés Fernández Díaz). Madrid, DELTA Publicaciones, 2009.

[4] J. E. Match & J. M. Birch, *Guide to successful thesis and dissertation*, 4th Ed. New York: Marcel Dekker, 1987.

3. Internet

La World Wide Web nos provee una variedad de recursos que incluyen artículos de libros, revistas, periódicos, documentos de agencias privadas y gubernamentales, etc. Éstas referencias deben proveer al menos, el título del recurso, fecha de publicación o fecha de acceso, y la dirección (URL) del recurso en la Web. En la medida que sea posible, se debe proveer el autor del recurso. Ejemplos:

[5] Tao Pang (2006). *An Introduction to Computational Physics*, 2nd Edition [Online]. Disponible en <http://www.physics.unlv.edu/~pang/cp2.html>

[6] (2003, Sep.) *IEEE computer society press proceedings author guidelines* [Online]. Disponible en http://www.computer.org/portal/pages/cscps/cps/cps_forms.html

4. Hiperenlaces

Es recomendable, si los artículos o libros tienen Digital object identifier (doi) o dirección web, utilizar el paquete `hyperref` de \LaTeX para la realización de los hiperenlaces que aparecerán en la versión digital de la revista. Se recomienda revisar el documento *Como usar la clase de documento \LaTeX 'killkanaT.cls' de la revista Killkana Técnica* en la sección VIII. Por ejemplo:

[1] I. Waller & R. Kapral, "Spatial and temporal structure in systems of coupled nonlinear oscillators", *Phys. Rev. A*, vol. 30, no. 4, pp. 2047-2055, Oct. 1984.

Lo mismo se aplica para las citas electrónicas:

[6] (2003, Sep.) *IEEE computer society press proceedings author guidelines*. Disponible en http://www.computer.org/portal/pages/cscps/cps/cps_forms.html

IV. CITAS TEXTUALES

Generalmente en publicaciones de áreas técnicas y científicas no se cita textualmente. Sin embargo, en caso de cita textual en un párrafo, se debe colocar el mismo entre comillas dobles. \LaTeX crea comillas dobles empleando los símbolos `\`texto'`, lo que resulta en: "texto".

Cuando se cita textualmente una frase de más de 4 líneas, se debe escribir en un párrafo aparte, con una sangría adicional y entre comillas. En este caso \LaTeX dispone del entorno `quote` y `quotation`. El entorno `quote` se emplea para una cita de un sólo párrafo, en cambio `quotation` para citas de varios párrafos. Para mayor información revisar el manual *Cómo usar la clase de documento \LaTeX 'killkanaT.cls', de la revista Killkana Técnica* sección VI.D.

V. FIGURAS Y TABLAS

Figuras claras comunican las ideas más efectivamente que el texto, ya que nuestros cerebros procesan las imágenes mucho más rápido que las palabras.

Los títulos de las figuras deben estar bajo ellas, como pie. \LaTeX proporciona el entorno de elemento flotante `figure`

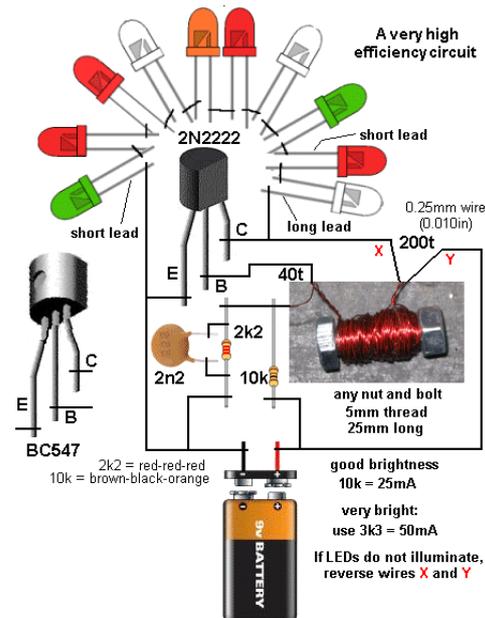


FIG. 1. Circuito de 10 leds con una batería de 9 V. Fuente <http://www.talkingelectronics.com>

TABLA I
FUNCIONES TRIGONÓMICAS PARA LOS ÁNGULOS NOTABLES

Ángulo	Seno	Coseno	Tangente
0°	0	1	0
30°	1/2	$\sqrt{3}/2$	$1/\sqrt{3}$
45°	$\sqrt{2}/2$	$\sqrt{2}/2$	1
60°	$\sqrt{3}/2$	1/2	$\sqrt{3}$
90°	1	0	—

que contienen el comando `caption` encargado de colocar la leyenda de la figura y las enumera automáticamente. Emplee la palabra "Fig.", dentro del texto, para referirse a las figuras. La Fig. 1 representa un ejemplo de una figura en un artículo.

No emplee imágenes borrosas y de mala calidad, es preferible utilizar archivos en formato vectorial, en nuestro caso pdf.

Para las tablas en cambio el título debe estar antes de la misma, como encabezado. \LaTeX proporciona el entorno de elemento flotante `table`, que también posee el comando `caption`, como en el caso de `figure`. La tabla I muestra un ejemplo realizado en \LaTeX de este tipo de elemento flotante.

Evite colocar figuras y tablas antes de su primera mención en el texto, a no ser que resulte muy conveniente por necesidades de maquetación y siempre que la primera mención quede muy próxima a la figura o tabla referida y en la misma página. Emplee el comando `label` para etiquetar las figuras y tablas para una fácil referencia dentro del texto con `ref`.

Si la tabla es muy compleja para realizarla en \LaTeX ,

realícelas en otro editor de tablas o texto y transfórmelas en imágenes de alta calidad de resolución. Insértelas en el documento como imágenes pero utilice el entorno flotante `table` para que sea considerada como una tabla.

Evite repetir contenidos, si coloca una imagen sobre la gráfica de unas variables pertenecientes a una tabla, no coloque la tabla. Prefiera las gráficas a las tablas.

Se recomienda leer la sección VI.A del documento *Como usar la clase de documento L^AT_EX 'killkana.cls' de la revista Killkana Técnica* en relación a todo lo expuesto anteriormente.

VI. EXPRESIONES MATEMÁTICAS

Todas las ecuaciones deben ir centradas y numeradas consecutivamente con números arábigos colocados entre paréntesis alineados a la derecha. Para esto se dispone principalmente del entorno de L^AT_EX `equation`, el cual enumera las ecuaciones automáticamente. La alineación de las ecuaciones y el formato de la numeración está controlado por la clase y el paquete `AMSMath`, por tanto, el autor puede utilizar todos los entornos de ecuaciones y definiciones proporcionadas por este paquete. A continuación un ejemplo de ecuación elaborada con el entorno `equation`:

$$\sigma = \sqrt{\frac{\sum_i^N (x_i - \bar{x})^2}{N - 1}}. \quad (1)$$

Las ecuaciones forman parte del texto del artículo y deben ser consideradas como una línea aparte en un párrafo, por lo tanto deben finalizar con algún signo de puntuación. Ejemplo:

La ecuación de la segunda ley de Newton tiene la forma:

$$F = m \frac{d^2x}{dt^2}, \quad (2)$$

donde m representa la masa y d^2x/dt^2 la aceleración del cuerpo.

Observe en el ejemplo anterior la coma al final de la ecuación, así mismo la siguiente línea comienza en minúscula y sin sangría.

VII. ABREVIATURAS Y ACRÓNIMOS

Defina las abreviaturas y los acrónimos la primera vez que aparezcan en el texto. Las abreviaturas muy comunes como IEEE, MKSA, c.c., c.a., no deben definirse. No emplee abreviaturas en el título del artículo, ni en el resumen a no ser que sean inevitables y muy comunes.

Se recomienda no abusar del empleo de abreviaturas y acrónimos para no aumentar el esfuerzo de lectura requerido por el artículo. Es preferible reducir su uso a casos muy conocidos incluso por los no especialistas, como c.c. (por corriente continua), c.a. (por corriente alterna), etc. No importa que de ello se derive cierto alargamiento del artículo.

VIII. AGRADECIMIENTOS

Esta sección es opcional. En este apartado se agradecen a las instituciones que colaboraron económicamente o con soporte de equipos, infraestructuras, etc. También se puede agradecer a personas que realizaron algún tipo de aportes a la investigación. El siguiente es un ejemplo de agradecimiento:

Los autores agradecen a la Dirección de Investigación de la Universidad Católica de Cuenca por el apoyo suministrado bajo el proyecto número XXX-X.

IX. ESPECIFICACIONES TÉCNICAS GENERALES

Debido a que la revista posee su propia clase de documento en L^AT_EX, el autor no debe preocuparse del tipo de letra, el tamaño de la misma, los márgenes de la página, las fuentes y tamaño para títulos, secciones, sub-secciones, etc. Los autores sólo deben enfocarse en escribir su trabajo, llenando los apartados correspondientes en la plantilla `killkanaT_template.tex` suministrada.

X. ACUERDO DE ORIGINALIDAD

Al enviar un artículo a la revista KILLKANA TÉCNICA el autor confirma que el trabajo entregado es inédito y original, que el mismo no ha sido publicado anteriormente en forma impresa o electrónica. A su vez se compromete a no enviar ni publicar en ningún otro medio antes de conocer el resultado de la evaluación por parte de la revista.

XI. NOTA PARA LOS AUTORES

Para finalizar se debe mencionar que “la principal limitante para la difusión y aplicación de los resultados de las investigaciones es precisamente la falta de habilidades para la redacción de artículo científico” [3], sin embargo, una vez que el autor se inicia en la tarea de escribir y redactar trabajos de investigación, descubre que la estructura de un artículo científico es relativamente fácil, ya que se cumplen con reglas específicas y técnicas que simplifican la tarea.

REFERENCIAS

- [1] R. Hernández, C. Fernandez, y P. Baptista, P., *Metodología de la Investigación*, 6ta. Ed. México: Mc Graw Hill Education, 2014
- [2] H. L. Swinney (2005), *Tips on Writing Papers* (actualizado en julio de 2014). Documento pdf disponible en <http://chaos.utexas.edu/wp-uploads/2014/07/tips-on-writing.pdf>
- [3] A. M. Contreras, & R. J. Ochoa, *Manual de redacción científica*. México: Ediciones de La Noche, 2010.

Recibido: 01 de enero de 2018

Aceptado: 01 de enero de 2018

Nombre1 y Apellido1: Breve resumen curricular del autor 1, no debe exceder de cuatro líneas

Nombre2 y Apellido2: Breve resumen curricular del autor 2, no debe exceder de cuatro líneas. nombre_autor1@correo.com



Cómo usar la clase de documento L^AT_EX ‘killkanaT.cls’ de la revista Killkana Técnica

How to use the L^AT_EX documentclass ‘killkanaT.cls’ of the Killkana Técnica journal

O. Alvarez-Llamoza

Centro de Investigación de Ingeniería, Industria, Construcción y TIC
Universidad Católica de Cuenca, Cuenca Ecuador
oalvarezll@ucacue.edu.ec

Resumen

Este artículo describe como usar la clase de L^AT_EX killkanaT.cls para producir documentos de alta calidad tipográfica, aptos para ser sometidos a la Revista Killkana Técnica de la Universidad Católica de Cuenca, Ecuador. Se proporciona al autor los conocimientos básicos para elaborar artículos empleando el sistema L^AT_EX y la clase suministrada. El propósito final es acelerar el proceso de edición y publicación en la revista facilitando las herramientas necesarias para producir artículos de excelente calidad.

Palabras clave: Artículo de investigación, clase de L^AT_EX, estilo, Killkana Técnica.

Abstract

This article describes how to use the killkana.cls’ LATEX type to produce high typesetting quality documents, suitable for submission to the Killkana Técnica Journal of “Universidad Católica de Cuenca”, Ecuador. The author is provided with the basic knowledge to write articles using the LATEX system and the given document type. The final purpose is to accelerate the publishing and editing process of the journal providing the necessary tools to produce articles of excellent quality.

Key words: Research article, L^AT_EX class, style, Killkana Técnica.

I. INTRODUCCIÓN

TENIENDO a disposición en su computadora el archivo de la clase `killkanaT.cls`, el de estilo `killkanaT.sty`, la plantilla `killkanaT_template.tex`, una distribución de L^AT_EX y conocimientos básicos sobre la edición de textos en dicho sistema, el autor puede crear artículos de investigación con una calidad profesional, de una forma rápida, cómoda, estructurada, y con un esfuerzo reducido en cuanto a la composición del estilo. El propósito principal de este documento es servir a los autores como manual de uso de la clase `killkanaT.cls`, en él se explican sus características y su comportamiento para la composición de artículos, de acuerdo a la normativa y el estilo establecido para la publicación en la revista KILLKANA TÉCNICA. Se recomienda leer la guía para la preparación de artículos *Instructivo para la preparación de manuscritos en la revista Killkana Técnica*, que se encuentra en un capítulo de éste ejemplar y en la página web de la revista <http://killkana.ucacue.edu.ec>, para las instrucciones, normas y como referencia en general para la

buen escritura de artículos de investigación.

Los archivos `killkanaT.cls` y `killkanaT.sty` se distribuyen a través de la licencia LPPL (L^AT_EX Project Public License) y pueden obtenerse de manera gratuita a través del contacto con el editor de la revista mediante el correo electrónico: killkana.editorial@ucacue.edu.ec o en la página web de la revista <http://killkana.ucacue.edu.ec>. Adicional a este documento también se distribuye el archivo `KillkanaT_template.tex`, con el propósito es servir como plantilla de documento, lista para su llenado y compilación.

Se asume que el lector de este artículo posee conocimientos básicos en el uso de L^AT_EX, se recomienda tener a mano la documentación que describa el funcionamiento básico de la composición de textos con este sistema. Libros como el de Borbón y Mora [1] y el de Tobias Oetiker [2] se pueden descargar libremente a través de la Internet.

Este artículo de información ha sido creado utilizando la clase `KillkanaT.cls` conjuntamente con el archivo de estilo `killkanaT.sty` y compilado con `pdflatex`. Dependiendo del sistema operativo se pueden emplear dis-

tintas distribuciones de \LaTeX y editores, los cuales pueden descargarse de manera gratuita:

- Para Windows se recomienda el compilador de \LaTeX MiKTeX que se puede descargar en la página <http://miktex.org/>. Para editar el documento puede utilizar TeXstudio (<http://www.texstudio.org/>) o Texmaker (<http://www.xmlmath.net/texmaker/>) entre otros.
- Para Mac OS se recomienda MacTeX (<https://tug.org/mactex/>), el cuál incluye el editor de \LaTeX Tex Shop (<http://pages.uoregon.edu/koch/texshop/texshop.html>). También puede emplear como editor TeXstudio o Texmaker, los cuales también tienen instaladores para Mac.
- Para Linux instalar la distribución TeX Live (<https://www.tug.org/texlive/>) y como editor se puede emplear Kile (<http://kile.sourceforge.net/>), TeXstudio y Texmaker.

Para someter un artículo de investigación al proceso de arbitraje de la revista el autor debe enviar al editor una carpeta del documento, con el archivo `.tex` (`killkanaT_template.tex` editado y con el nombre modificado) y las correspondientes figuras en los formatos adecuados para su inserción y compilación en `pdflatex`. El editor, en un lapso de tiempo acordado o el que la normativa lo sugieran, revisará el documento con el objeto de poder realizar correcciones de formato y editar las definiciones que corresponden al editor, descritas en la sección **XI**.

II. PAQUETES DE \LaTeX REQUERIDOS

`killkanaT.cls` requiere el uso de varios paquetes de \LaTeX , los cuales permiten hacer la composición de los artículos de una forma elegante y eficiente, puesto que están definidos para solucionar tareas específicas. Estos paquetes habitualmente se encuentran incluidos de forma estándar en las distribuciones básicas más conocidas de \LaTeX , y si no, se debe conocer los procedimientos a seguir para la instalación de nuevos paquetes, en el caso de que la distribución en uso no contenga por defecto a alguno de ellos. Los siguientes paquetes, con las opciones entre corchetes, son requeridos para la compilación de los artículos de KILLKANA TÉCNICA: `Fancyhdr`, `Float`, `Babel[spanish,USenglish]`, `Inputenc[utf8]`, `AMSMath`, `Xcolor[pdftex,usenames,dvipsnames,svgnames,table]`, `Graphicx` y `Hyperref[draft,colorlinks]`. Los paquetes anteriores se encuentran especificados en el archivo de estilo `killkanaT.sty` proporcionado.

Si se requiere enviar opciones adicionales a los paquetes definidos se debe hacer uso del comando

```
\PassOptionsToPackage{<opciones>}
{<paquete>}
```

antes del comando `\begin{document}` que inicia el código fuente del documento del artículo (`archivo.tex`).

Adicionalmente, si requiere el uso de otro paquete para su artículo lo puede definir también en el preámbulo del documento antes señalado (`archivo.tex`) empleando el comando

```
\usepackage[opciones]{paquete}
```

III. OPCIONES PARA EL IDIOMA

La REVISTA KILLKANA TÉCNICA acepta trabajos en español y en inglés. El código de los documentos compuestos con \LaTeX debe contener la definición del idioma del documento que se va a componer. Para esto se tiene el comando `\documentclass[<opciones>]{<clase>}`, que permite el envío de opciones a la clase y define el idioma del documento y cambian su comportamiento. La clase `killkanaT.cls` tiene sólo dos opciones: `spanish` (por omisión) y `english`, ambas corresponden al idioma que será utilizado por el paquete Babel como el idioma principal en que será construido el artículo.

A. Artículo en español: *spanish*

La siguiente es la opción por omisión para un artículo, no es necesaria su declaración: selecciona la opción `spanish` del paquete Babel empleando el comando `\selectlanguage{spanish}`.

```
\documentclass[spanish]{killkanaT}
```

B. Artículo en inglés: *english*

Esta opción se debe utilizar cuando el documento tiene como idioma principal el inglés: selecciona la opción `USenglish` del paquete Babel con el comando `\selectlanguage{USenglish}`. Así, se debe colocar:

```
\documentclass[english]{killkanaT}
```

El estilo `killkanaT.cls` se encarga de colocar correctamente los lenguajes en las leyendas del documento donde se cambia del lenguaje principal al otro; figuras, tablas, referencias, etc. Sin embargo es obligatorio la inclusión de títulos, resúmenes y palabras clave en idioma español e inglés. Sólo es necesaria la inclusión de una opción de idioma si va a escribir su trabajo en inglés.

IV. LA PRIMERA PÁGINA DEL ARTÍCULO

La primera parte del documento que corresponde a la página del título posee adicionalmente el nombre del autor, la institución de afiliación, el correo electrónico, el resumen y las palabras clave. Estos elementos deben ser construidos utilizando el comando estándar `maketitle` dentro del cuerpo del documento. Los elementos de la sección del título deben ser declarados antes del uso de este comando, estos elementos se definen a continuación.

A. Título

Un artículo debe contener un título principal en el idioma “principal” del documento (español), y un título en el idioma secundario (inglés). Para definir el título se dispone del comando:

```
\title[<corto>]{<principal>}{<secundario>}
```

Este comando define el título en el idioma principal del documento en el parámetro `principal`, el título del idioma secundario en el parámetro `secundario` y un título opcional (`corto`), visible en los encabezados de páginas impares del documento. En el parámetro opcional `corto`, utilizable sólo en caso de que el título principal sea muy largo y se desborde de la línea del encabezado. En tal caso, el autor debe utilizar el mismo título principal en el parámetro opcional `corto` interrumpido con tres puntos suspensivos (`\ldots`) donde crea conveniente.

Por defecto, el título principal debe ser en idioma español y el título secundario en inglés, lo contrario debe cumplirse cuando se utiliza la opción `english` en la declaración `documentclass` (sección III.).

B. Autor(es), afiliación(es) y correo autor principal

La declaración del nombre del autor, la institución y el correo en un artículo se hacen con los comandos:

```
\author{<autor>}
\affiliation{<institución>}
\email{<correo electrónico>}
```

El nombre del autor sigue la normas del *Institute of Electrical and Electronics Engineers* (IEEE). En los campos requeridos se debe colocar lo siguiente:

- Identificación del autor: Nombre y Apellido.
- La afiliación que debe escribirse de forma ascendente respecto a la importancia del organismo que se mencione. Se debe introducir saltos de línea (`\\`) para controlar la forma de la afiliación, se recomienda el uso de una línea para la afiliación académica en sí y otra contigua para su lugar de origen.
- El correo electrónico del autor o en el caso de varios autores, sólo del autor principal.

Así, llenando los campos indicados se tiene el ejemplo:

```
\author{Isaac Newton}
\affiliation{Departamento de Física,
Universidad de Cambridge\\
Cambridg CB2 1TN, Reino Unido}
\email{inewton@ocambridge.edu.euk}
```

Lo cual produce:

Isaac Newton
Departamento de Física, Universidad de Cambridge
Cambridg CB2 1TN, Reino Unido
inewton@ucambridge.edu.uk

Para varios autores los nombres se deben separar por comas y de ser necesario colocar sólo la inicial de los nombres. Si se tienen afiliaciones diferentes para los distintos autores, éstas deben estar definidas dentro del comando `affiliation` seguida una de otra y precedidas en su

primera línea de una numeración, para lo que se dispone del comando `\affil{<Z>}`, dónde `Z` es un número natural arábigo. Este comando debe colocarse a su vez luego del nombre de cada autor al que le corresponda la afiliación en el comando `author`.

Si se tiene más de un autor se debe incluir antes de la dirección de correo electrónico y luego del nombre del autor principal el símbolo “*” que indica la correspondencia entre el autor y su correo electrónico. A pesar de que se asume que el autor principal es el primero a quien se le hace mención, sin embargo puede modificarse el orden. Un ejemplo completo de un apartado de autores en la página de título se muestra como sigue:

```
\author{I. Newton\affil{1}*,
E. Halley\affil{2} & G. Leibniz\affil{3}}
\affiliation{
\affil{1}Departamento de Física,
Universidad de Cambridge\\
Cambridg CB2 1TN, Reino Unido\\
\affil{2}
Grupo de Astronomía, Universidad de Oxford\\
Oxford, Reino Unido \\
\affil{3} Departamento de Matemáticas,
Universidad de Altdorf\\
Altdorf bei Nurnberg, Nuremberg, Alemania
}
\email{*inewton@ocambridge.edu.euk}
```

Lo cual produce:

I. Newton¹*, E. Halley² & G. Leibniz³
¹Departamento de Física, Universidad de Cambridge
Cambridg CB2 1TN, Reino Unido
²Grupo de Astronomía, Universidad de Oxford
Oxford, Reino Unido
³Departamento de Matemáticas, Universidad de Altdorf
Altdorf bei Nurnberg, Nuremberg, Alemania
*inewton@ucambridge.edu.uk

Si se requiere asignar dos afiliaciones a un autor, se separan los números correspondientes con un guion en el comando `affil`, por ejemplo: `\affil{1-3}`.

El autor obligatoriamente debe seguir el formato aquí descrito con el objeto de mantener consistencia en toda la forma de la publicación. Los elementos antes descritos deben ser colocados antes del cuerpo del documento (`\begin{document}`) en el preámbulo del mismo.

V. RESUMEN Y PALABRAS CLAVE

Para este elemento de la página de título se dispone del comando `\spanishabstract{<resumen>}` para el resumen en español y `\englishabstract{<abstract>}` para el resumen en inglés, estos comandos deben ser declarados luego del inicio del cuerpo del documento (`\begin{document}`) y necesariamente antes del comando `maketitle`. El resumen debe tener un máximo de doscientas cincuenta (250) palabras para todos artículos. Revisar el documento *Instructivo para la preparación de manuscritos en la revista Killkana Técnica*, sección III.B, para mayor información.

Para las palabras clave el autor debe definir de tres (3) a seis (6) palabras relacionadas con la naturaleza del artículo. Los comandos `\spanishkeywords{<palabras>}` para las palabras clave en español y `\englishkeywords{<words>}` para las palabras clave en inglés. En los argumentos de ambos comandos se debe colocar las palabras clave separadas por comas y sin el punto al final. Así tenemos el ejemplo:

```
\spanishabstract{Éste es un ejemplo de
resumen en español...}

\spanishkeywords{primera, segunda, tercera}

\englishabstract{This is an example of
abstract in english...}

\englishkeywords{first, second, third}
```

Lo cual produce:

Resumen

Éste es un ejemplo de resumen en español...

Palabras clave: primera, segunda, tercera.

Abstract

This is an example of abstract in english...

Key words: first, second, third.

VI. CUERPO DEL ARTÍCULO

La creación de los apartados en los artículos creados con esta clase se realiza igual que la forma estándar de \LaTeX . Se dispone de comandos para los títulos de las secciones, sub-secciones y sub-sub-secciones que requieran para componer el trabajo. La aplicación de comandos para estas divisiones del documento es de la forma:

```
\section{Título de la sección}
Contenido de la sección ...

\subsection{Título de la subsección}
Contenido de la sub-sección ...

\subsubsection{Título de la subsection}
Contenido de la sub-sub-section ...
```

Lo cual resulta:

I. TÍTULO DE LA SECCIÓN

Contenido de la sección ...

A. Título de la sub-sección

Contenido de la sub-sección ...

2. Título de la sub-sub-sección

Contenido de la sub-sub-sección ...

A. Elementos flotantes

`killkanaT.cls` emplea el paquete `Float` que permite un mejor control sobre el posicionamiento de los elementos flotantes. Se identifican como elementos flotantes a las ecuaciones, las figuras y las tablas, porque son

componentes del documento que se posicionan fuera de los párrafos, como si estuvieran “flotando” sobre la página.

1. Ecuaciones

Todas las ecuaciones deben ir enumeradas consecutivamente con números arábigos colocados entre paréntesis alineados a la derecha. Para esto \LaTeX dispone principalmente del entorno `equation`. La alineación de las ecuaciones, la numeración y su formato está controlado automáticamente por la clase y el paquete `AMSMath`, lo que brinda al autor todos los entornos de ecuaciones y definiciones proporcionadas por este paquete.

A continuación el ejemplo de ecuación o fórmula en código \LaTeX :

```
\begin{equation}
\sigma = \sqrt{\frac{\sum_{i=1}^N (x_i - \overline{x})^2}{N-1}}
\end{equation}
\label{desv-est}
```

Lo cual produce la ecuación:

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N - 1}}. \quad (1)$$

El autor debe tener en cuenta que en el código no debe haber salto de línea, o una línea en blanco, entre el entorno matemático y el párrafo que lo precede, de lo contrario ocurrirá que en el documento compilado aparezca un espacio mayor antes de la ecuación que después de ella.

Si al final de la ecuación va una coma, significa que la siguiente línea no tiene sangría y esto se logra no dejando espacio entre la ecuación y el siguiente párrafo. Recuerde que una ecuación es parte del texto y debe finalizar con algún signo de puntuación. Revisar el documento *Instructivo para la preparación de manuscritos en la revista Killkana Técnica*, sección IV.

Si una ecuación es muy larga, se puede hacer uso de entornos como `multline` del paquete `AMSMath` u otros entornos que permitan la manipulación adecuada de una ecuación en distintas líneas.

Finalmente el comando `\label{<etiqueta>}` sirve para identificar la ecuación y referenciarla mediante el comando `\ref{<etiqueta>}` en alguna parte del artículo. Por ejemplo, tomando la etiqueta dada a la ecuación anterior tenemos:

```
La ecuación (\ref{desv-est}) es la desviación
estándar ...
```

Lo que produce:

La ecuación (1) es la desviación estándar ...

2. Tablas

\LaTeX dispone para la inclusión de tablas del entorno para elementos flotantes `table` y del de construcción de tablas `tabular`.

```
\begin{table}[<posición>
\centering
\caption{Esta tabla muestra algunos datos.}
\begin{tabular}{|l|}
... Contenido tabla ...
\end{tabular}
\label{<etiqueta>}
\end{table}
```

donde las opciones de “posición” son: h (here) aquí, t (top) arriba de la columna, b (bottom) al final de la columna, y p (page) página completa.

El comando `caption` se ubica antes del entorno `tabular`, porque el estilo establece colocar la leyenda de la tabla antes de la tabla misma. Por ejemplo:

```
\begin{table}[h]
\centering
\caption{Funciones seno y coseno para los ángulos notables}
\begin{tabular}{c|c|c}
Angulo & Seno & Coseno & \\
\hline
0\degree & $0$ & $1$ & \\
30\degree & $1/2$ & $\sqrt{3}/2$ & \\
45\degree & $\sqrt{2}/2$ & $\sqrt{2}/2$ & \\
60\degree & $\sqrt{3}/2$ & $1/2$ & \\
90\degree & $1$ & $0$ & \\
\end{tabular}
\label{funciones}
\end{table}
```

Lo anterior da como resultado:

Ángulo	Seno	Coseno
0°	0	1
30°	1/2	$\sqrt{3}/2$
45°	$\sqrt{2}/2$	$\sqrt{2}/2$
60°	$\sqrt{3}/2$	1/2
90°	1	0

Si la tabla fue realizada con otro software, se puede exportar como una figura y ser insertada con el comando `includegraphics`, como se verá en la siguiente subsección.

3. Figuras

Las fotografías, diagramas, y gráficos en KILLKANA TÉCNICA se colocan empleando el entorno de flotantes `figure` y el comando de inserción de gráficos `includegraphics` de \LaTeX . Éste último forma parte del paquete `graphicx`, ya incluido en el archivo de estilo `killkanaT.sty`.

Para las imágenes a insertar se recomienda emplear formato de archivo vectorial, en nuestro caso formato `pdf`. No obstante, formatos como `jpeg` y `png` son efectivamente compilados por `pdflatex`, con los cuales la calidad de impresión es buena, pero las imágenes en el documento digital en algunos casos pueden carecer de nitidez, sobre todo cuando se amplían. Para la edición en línea de la revista se aceptan figuras a colores, sin embargo, se debe considerar que para la versión impresa todas las figuras deben tener

una coloración en escala de grises, bien contrastadas. Un ejemplo de inclusión de figura es el siguiente:

```
\begin{figure}[h]
\centering
\includegraphics[scale=0.45]{leds}
\caption{Circuito de 10 leds con una batería de 9 V. Fuente
\url{http://www.talkingelectronics.com}}
\label{fig:leds}
\end{figure}
```

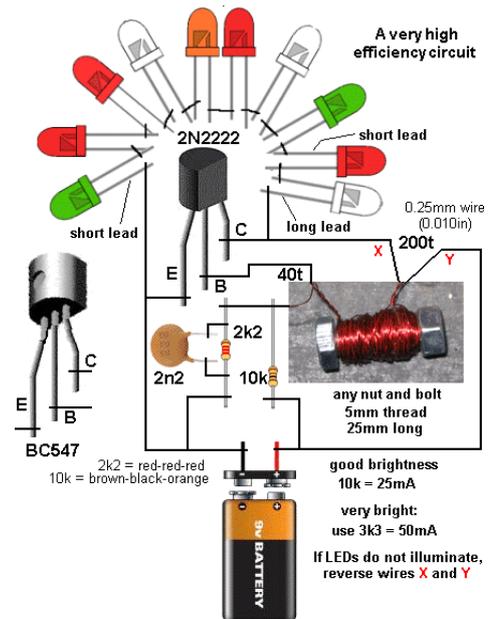


FIG. 1. Circuito de 10 leds con una batería de 9 V. Fuente <http://www.talkingelectronics.com>

La versión de la fotografía 1 para la edición impresa da como resultado la Fig. 2, que se muestra en la siguiente página.

El comando `url` utilizado en el `caption` de la figura se emplea para realizar enlaces con páginas web, revisar la sección VIII.

4. Tablas insertadas como figuras

Dependiendo del caso, el autor puede tener ya elaborada una tabla con otro entorno informático, en tal caso puede exportarla como imagen y adjuntarla empleando el entorno `table`, pero insertándola como imagen utilizando el comando `includegraphics`. Ejemplo:

```
\begin{table}[h]
\centering
\caption{Prefijo para múltiplos de bits o bytes.}
\includegraphics[scale=0.45]{tabla-pref.png}
\label{prefijos}
\end{table}
```

Lo que resulta después de compilado la tabla I de la siguiente página.

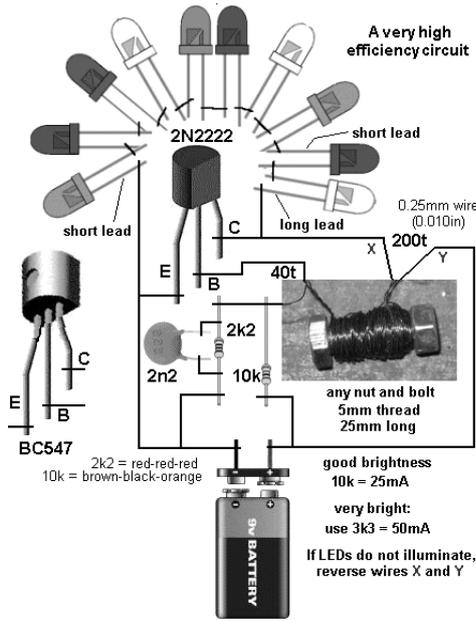


FIG. 2. Circuito de 10 leds con una batería de 9 V. Fuente <http://www.talkingelectronics.com>

TABLA I
PREFIJO PARA MÚLTIPLOS DE BITS O BYTES.

Decimal		Binary		
Value	SI	Value	IEC	JEDEC
1000	k kilo	1024	Ki kibi	K kilo
1000 ²	M mega	1024 ²	Mi mebi	M mega
1000 ³	G giga	1024 ³	Gi gibi	G giga
1000 ⁴	T tera	1024 ⁴	Ti tebi	-
1000 ⁵	P peta	1024 ⁵	Pi pebi	-
1000 ⁶	E exa	1024 ⁶	Ei exbi	-
1000 ⁷	Z zetta	1024 ⁷	Zi zebi	-
1000 ⁸	Y yotta	1024 ⁸	Yi yobi	-

5. Flotantes que abarcan 2 columnas

En caso de que se requiera la inclusión de figuras y tablas que deban ocupar el espacio correspondiente a las dos columnas en que está estructurado el texto, se puede hacer uso de los entornos de \LaTeX `figure*` y `table*`. Éstos entornos tienen ciertas limitantes y carecen de una completa libertad de posicionamiento, por tanto hay que tener en cuenta que por omisión los elementos serán colocados una página después de donde son llamados y en la parte superior de la misma. Se puede hacer uso del paquete `dblfloatfix` para extender las prestaciones de \LaTeX y brindar estabilidad en el manejo de elementos flotantes en este caso, también el uso de este paquete permite la

colocación de elementos en la parte inferior de la página (opción 'b') y la inserción de elementos simultáneos. Si se requiere que una ecuación ocupe las dos columnas se puede hacer uso del entorno `equation`, u otro entorno matemático, dentro de un entorno `figure*` o `table*`.

B. Entornos de listas

Para la creación de listas se utilizan los entornos \LaTeX : `enumerate`, `itemize` y `description`. Ejemplo:

```
\begin{itemize}
  \item Triángulo equilátero. Posee tres lados iguales y tres ángulos iguales de 60\degree.
  \item Triángulo isósceles. Posee dos lados y dos ángulos iguales.
  \item Triángulo escaleno. No hay lados ni ángulos iguales.
\end{itemize}
```

- Lo cual resulta en la lista:
- Triángulo equilátero. Posee tres lados iguales y tres ángulos iguales de 60°.
 - Triángulo isósceles. Posee dos lados y dos ángulos iguales.
 - Triángulo escaleno. No hay lados ni ángulos iguales.

Con `enumerate` se obtendría una lista numerada.

C. Notas al pie de página

Para agregar notas al pie de página el autor debe utilizar el comando de \LaTeX `\footnote{<nota>}`, el cual las numera automáticamente y las coloca al final de la columna. Ejemplo:

```
En matemáticas los cuaterniones\footnote{Los cuaterniones fueron descritos por primera vez por el físico y matemático irlandés William Rowan Hamilton en 1843.} son un sistema de números que extienden a los números complejo.
```

Lo que resulta:

- En matemáticas los cuaterniones^a son un sistema de números que extienden a los números complejo.
- ^aLos cuaterniones fueron descritos por primera vez por el físico y matemático irlandés William Rowan Hamilton en 1843.

D. Citas textuales

Generalmente en publicaciones de áreas técnicas y científicas no se cita textualmente. Sin embargo, en caso de cita textual en un párrafo, se debe colocar el mismo entre comillas dobles. \LaTeX crea comillas dobles empleando los símbolos ```texto''`, lo que resulta en: "texto".

Cuando se cita textualmente una frase de más de 4 líneas, se debe escribir en un párrafo aparte, con una sangría adicional y entre comillas. En este caso se dispone del entorno `quote` y `quotation`. El entorno `quote` se emplea para una cita de un sólo párrafo, en cambio `quotation` para citas de varios párrafos. Como ejemplo de uso del entorno `quote`, empleando la referencia definida anteriormente, se tiene:

```
\begin{quote}
  ``Un planteamiento cualitativo es como
  ``ingresar a un laberinto''. Sabemos dónde
  comenzamos, pero no dónde habremos de terminar.
  Entramos con convicción, pero sin un mapa
  detallado, preciso. Y de algo tenemos certeza:
  deberemos mantener la mente abierta y estar
  preparados para improvisar''
  \cite{Hernandez2014}
\end{quote}
```

Lo que resulta

“Un planteamiento cualitativo es como “ingresar a un laberinto”. Sabemos dónde comenzamos, pero no dónde habremos de terminar. Entramos con convicción, pero sin un mapa detallado, preciso. Y de algo tenemos certeza: deberemos mantener la mente abierta y estar preparados para improvisar”
[1]

Nota: en el ejemplo anterior se ha omitido la caja coloreada que caracteriza en este documento a los productos obtenidos después de la compilación, para que el autor aprecie directamente el resultado del entorno `quote` en un documento.

VII. REFERENCIAS BIBLIOGRÁFICAS

Las referencias bibliográficas deben aparecer en una sección al final del artículo y deben estar enumeradas según el orden de aparición en el documento y siguiendo las normas del IEEE. Los elementos que las conforman y el orden depende del tipo de documento. Por ejemplo para artículos tenemos:

- 1) Iniciales y Apellido del autor.
- 2) “Título del artículo entre comillas”.
- 3) *Título abreviado de la revista en letra cursiva.*
- 4) Volumen abreviado vol.
- 5) Número abreviado no.
- 6) Páginas abreviado pp.
- 7) Mes, año.

En caso de sitios Web: Iniciales y apellidos del autor (si lo tuviere). (año, mes, día). Título [Tipo de soporte / medio]. Volumen (número) (si lo tuviere). Disponible en : <http://...>

En la Internet se encuentra suficiente información sobre el estilo del IEEE, documentos en pdf como [Estilo de Referencias IEEE](#) de la Facultad de Derecho de la UNAM o el [FORMATO IEEE: Estilo y referencias bibliográficas](#) de la Universidad de Málaga, pueden consultarse.

A. Forma manual

La lista de referencias y la citas bibliográficas se pueden realizar de forma manual con el empleo del entorno `thebibliography`. El entorno `thebibliography` se coloca al final del documento y antes de `\end{document}`:

```
\begin{thebibliography}{9}
  \bibitem{etiqueta1} Contenido de la referencia
  \bibitem{etiqueta2} Contenido de la referencia
\end{thebibliography}
```

El elemento dentro de las llaves “{ }” corresponde a la etiqueta que permite ejecutar el comando para la cita. Por ejemplo:

```
\bibitem{Hernandez2014} R. Hernández, C.
  Fernandez, y P. Baptista, {\it Metodología
  de la Investigación}, 6ta. Ed. México:
  Editorial Mc Graw Hill Education, 2014.
```

Lo cual da como resultado en la sección “Referencias”:

[1] R. Hernández, C. Fernandez, y P. Baptista, *Metodología de la Investigación*, 6ta. Ed. México: Editorial Mc Graw Hill Education, 2014.

La etiqueta de la referencia se utiliza para realizar la cita bibliográfica en el texto. El autor debe utilizar el comando `\cite{<etiqueta>}` propio de \LaTeX , el cual numera automáticamente según el orden escrito en `thebibliography` a las referencias. Así, siguiendo el ejemplo anterior, se citaría en un párrafo de la siguiente forma:

```
De la metodología para realizar un trabajo de
investigación \cite{Hernandez2014} se puede
concluir...
```

Dando como resultado:

De la metodología para realizar un trabajo de investigación [1] se puede concluir...

B. Con BibTeX

Se puede utilizar BibTEX para componer la lista de referencias, para lo cual se necesita un archivo de bases bibliográficas `.bib` y especificar el estilo. Ésta información se debe colocar al final del documento, en lugar del entorno `thebibliography` y antes de `\end{document}`, empleando los comandos:

```
\bibliography{<file>}
\bibliographystyle{<style>}
```

donde el argumento `file` es el nombre del archivo `.bib` de la base bibliográfica (no escribirlo con la terminación `.bib`) y `style` es el estilo para las referencias, en el caso de KILLKANA TÉCNICA colocamos `ieeetr`. Al compilar se generará un archivo con terminación `.bbl`, el cuál guarda la información de las citas tal como si las hubiéramos escrito manualmente dentro del entorno `thebibliography`.

Al finalizar el documento el autor debe cambiar de nombre del archivo `.bbl`, por ejemplo llamarlo `Referencias.tex`. Por último sustituir los comandos `bibliography` y `bibliographystyle` por `\input{Referencias}`. El archivo de referencias (en el ejemplo `Referencias.tex`), conjuntamente con su archivo de artículo `.tex` y las figuras se deben enviar al editor de la revista. No envíe su base de datos bibliográfica (archivo `.bib`).

VIII. HIPERENLACES

Es recomendable, si los artículos o libros tienen Digital object identifier (doi) o dirección web, uti-

lizar el paquete `hyperref` de \LaTeX para la realización de los hiperenlaces que aparecerán en la versión digital de la revista. El paquete `hyperref` está incluido en el archivo de estilo `killkanaT.sty`: `\usepackage[colorlinks]{hyperref}`. La opción `colorlinks` crea enlaces a color de las ecuaciones, figuras citas, páginas web, correos electrónicos, etc. Otra opción importante es `draft`, la cual inhabilita los hiperenlaces, lo cual es útil para la versión impresa.

`hyperref` crea automáticamente los enlaces a ecuaciones, figuras y citas dentro del documento, sin embargo para enlaces con el navegador a las páginas web y correos electrónicos se debe especificar su dirección en la Internet. Dentro del documento se tienen las siguientes opciones del uso de `hyperref`:

```
\url{<dirección de página web>}
\href{<dirección de página web>}{<texto que aparece en el documento>}
\href{mailto:<correo electrónico>}{<texto que aparece en el documento>}
```

Empleando el comando `url` anterior con el argumento requerido se un ejemplo de enlace para una página web en las referencias:

```
\bibitem{Oetiker2014} Tobias Oetiker. (2014). {\it A short introduction to \LaTeX2e} [online] (versión en español). Disponible en \url{https://www.ctan.org/tex-archive/info/lshort}.
```

Con lo cual se obtiene en las referencias:

[2] Tobias Oetiker (2014). *A short introduction to $\text{\LaTeX}2e$* [online] (versión en español). Disponible en <https://www.ctan.org/tex-archive/info/lshort/>.

Empleando el comando `href` y con los argumentos requeridos se da un ejemplo de enlace para un artículo:

```
\bibitem{Alvarez2008} O. Alvarez-Llamoza y M. G. Cosenza, "Generalized synchronization of chaos in autonomous systems", \href{http://link.aps.org/doi/10.1103/PhysRevE.78.046216}{{\it Phys. Rev. E}}, vol. 78, no 4, pp. 046216, Oct. 2008}.
```

Con lo cual se obtiene en las referencias:

[3] O. Alvarez-Llamoza y M. G. Cosenza, "Generalized synchronization of chaos in autonomous systems", *Phys. Rev. E*, vol. 78, no 4, pp. 046216, Oct. 2008.

IX. BIBLIOGRAFÍA DE LOS AUTORES

El documento debe contener una pequeña biografía de los autores, no más de cuatro líneas en letra cursiva, del autor o autores ubicada al final del texto del documento. La biografía de un autor debe contener al principio (sin sangría) el nombre del autor en negrita, en la forma Nombre Apellido, seguido por la breve descripción de su experiencia académica y/o profesional, de un máximo de 4 líneas. En caso de no ser el autor principal debe culminar con su dirección de correo electrónico.

La clase `killkanaT.cls` dispone del comando `\biography{<lista-biografías>}`, que debe ser declarado en el preámbulo del documento, para construir la biografía al final del texto del documento. El argumento corresponde a la biografía del o los autores, con el formato descrito anteriormente, donde para cada autor corresponde un párrafo específico. Dentro de `biography` se dispone del comando `\name{<Nombre Apellido>}{<biografía>}`, para definir la biografía de cada autor. De esta manera, una biografía puede declararse de la siguiente forma:

```
\biography{
\name{Pedro Pérez Ceráz} Ing. eléctrico con maestría en Energía. Profesor de la Universidad del Sur con más de 10 años de experiencia. Realiza investigaciones de generación eléctrica en sistemas híbridos.

\name{Ana Lisa Paredes}{Ing. Civil con PhD en Construcciones Industriales. Profesora de la Universidad del Norte con 6 años de experiencia docente y de investigación, Correo electrónico: anparedes@udelnorte.edu.zn}
}
```

Lo que genera al final del artículo:

Pedro Pérez Ceráz: *Ing. eléctrico con maestría en Energía. Profesor de la Universidad del Sur con más de 10 años de experiencia. Realiza investigaciones de generación eléctrica en sistemas híbridos.*

Ana Lisa Paredes: *Ing. Civil con PhD en Construcciones Industriales. Profesora de la Universidad del Norte con 6 años de experiencia docente y de investigación. Correo electrónico: anparedes@udelnorte.edu.zn*

X. ENCABEZADOS Y PIE DE PÁGINA

Los encabezados y pie de página varían según la página (impar o par). Entre los elementos que constituyen los encabezados el autor sólo necesita tener control sobre el nombre del autor o de los autores en las páginas pares y del título en las páginas impares. Para el primero se dispone del comando `\shortauthors{<nombre>}`, con el fin de definir el nombre del autor o los autores en los encabezados de las páginas pares. Debe escribirse de la forma: { Nombre Apellido } para un solo autor, { Apellido1 y Apellido2 } para 2 autores y, en el caso de haber más de un autor se escribe el apellido del autor principal seguido de las palabras "et al.", que significa "y otros".

Como se había mencionado anteriormente (sección IV.A.) el título corto se definirá como parámetro adicional del comando `title` para las páginas impares.

En la siguiente sección se describe el uso de comandos para el editor que permiten definir los otros elementos del encabezado.

XI. COMANDOS PARA EL EDITOR

En el preámbulo del documento se encuentran comentarios varios comandos que son para ser llenados por el editor, en ellos se especifica información sobre la revista

y el artículo recibido. Los comandos del editor proveen información sobre el tipo de artículo; volumen, número, año y meses correspondientes al ejemplar de la revista; ISSN impreso y electrónico de la revista; página inicial del artículo; y finalmente la fecha de recepción y fecha de aceptación del artículo.

XII. BALANCE DE COLUMNAS EN LA ÚLTIMA PÁGINA

La última página del documento debe tener un balance entre las columnas, de forma que la columna de la izquierda no sea más larga que la de la derecha o por lo menos no lo suficiente como para que se vea una página desequilibrada. Para esto el autor o el editor pueden controlar la longitud de ambas columnas añadiendo un salto manual en la columna de la izquierda donde lo vea necesario.

Para hacer el balance de forma manual se recomienda el uso del comando definido en la clase: `\columnbreak`, que obliga a un salto de columna, si dicho salto debe hacerse dentro de una línea de un párrafo. A su vez se puede utilizar el comando `\newpage` en caso de que se haga el

salto de columna justo al terminar un párrafo.

XIII. PÁGINA PAR AL FINAL DEL DOCUMENTO

La clase `killkanaT.cls` agrega una página en blanco al final del documento si éste termina en una página impar, esto con la finalidad de favorecer la composición del número de la revista impresa, que es una compilación de los artículos aceptados y en los cuales todos deben comenzar desde una página impar.

REFERENCIAS

- [1] A. Borbón & W. Mora (2014), *Edición de textos científicos L^AT_EX 2014* [online], Libro de la Revista digital Matemática, Educación e Internet, Instituto tecnológico de Costa Rica. Disponible en <http://tecdigital.tec.ac.cr/revistamatematica/Libros/index.htm>
- [2] Tobias Oetiker. (2014). *A short introduction to L^AT_EX 2_ε* [online] (versión en español). Disponible en <https://www.ctan.org/tex-archive/info/lshort/>.

Orlando Alvarez-Llamoza: *PhD. en Física Fundamental con 19 años de experiencia universitaria. Realiza investigaciones computacionales y teóricas en sistemas complejos, dinámica no lineal, física estadística, y aplicaciones interdisciplinarias.*

Sistema y proceso de arbitraje de la revista Killkana Técnica

System and process of refereeing in the Killkana Técnica journal

Consejo editorial - Revista Killkana Técnica
Universidad Católica de Cuenca, Cuenca Ecuador
killkana.editorial@ucacue.edu.ec

Resumen

La publicación de trabajos científicos en la revista Killkana Técnica se lleva a cabo a través del proceso de evaluación de dos revisores o pares ciegos, donde la identidad de los evaluadores y de los autores no son reveladas a ninguna de las partes. Este proceso colaborativo permite que los artículos recibidos sean valorados y comentados por expertos independientes a la institución, garantizando que los trabajos cumplan con un mínimo número de parámetros de calidad exigidos a nivel internacional. Adicionalmente se indican los factores que pueden influir en el tiempo del proceso de revisión y aceptación de un artículo.

Palabras clave: Proceso de publicación, revisión de pares ciegos, evaluadores, tiempo de espera.

Abstract

The publication of scientific papers in Killkana Técnica Journal is carried out through the evaluation process of two reviewers or blind peers, where the evaluators and the authors' identities are not exposed to any of the parties. This collaborative process allows the articles received to be evaluated and commented by experts who are independent to the institution, making sure that the research works comply with a minimum number of quality parameters required at the international level. Additionally, the factors that may influence the time of the review process and acceptance of the article are stated.

Key words: Publication processes, .

La aceptación de trabajos en la revista Killkana Técnica se lleva a cabo a través de la revisión de pares, también conocidos como referís. Éste proceso colaborativo permite que los manuscritos enviados a las revistas sean evaluados y comentados por expertos independientes a la institución, dentro de la misma área de investigación. La evaluación y crítica realizada por los pares evaluadores genera una retroalimentación para que el autor mejore y afine su trabajo, a la vez que permite al editor y/o equipo editorial valorar la calidad del artículo para ser publicado en la revista.

La revisión de trabajos realizado por pares expertos es un proceso formal y explícito para comunicaciones científicas, y se emplea desde las primeras revistas científicas aparecidas hace más de trescientos años.

El proceso de revisión de la revista Killkana Técnica es el de dos revisores o pares ciegos: la identidad de los referís y de los autores no son reveladas a ninguna de las partes.

I. BENEFICIOS DE LA REVISIÓN POR PARES

- El autor recibe una detallada y constructiva retroalimentación de parte de expertos en el área.
- El proceso puede alertar a los autores de errores o vacíos en la literatura que se pudieron haber omitir.
- Puede ayudar a que el artículo sea más accesible a los lectores de las revistas.

- Podría propiciar una discusión constructiva entre el autor, el referí y el editor, acerca de un campo o tópico científico.

II. LABOR DEL EDITOR O EQUIPO EDITORIAL

El equipo editorial, al que se denomina también “Editor”, está conformado por los Coordinadores de los Centros de Investigación de la Universidad Católica de Cuenca, el cual tiene las responsabilidades descritas a continuación:

- El editor considerará si un trabajo es lo suficientemente bueno para enviarlo o no a los pares evaluadores. En tal sentido debe revisar que el artículo se ajusta a objetivos y alcance, directrices y las instrucciones dadas a los autores de las revistas. Adicionalmente debe asegurarse que el contenido del manuscrito representa una contribución seria y significativa al campo del conocimiento.
- Los artículos pueden ser rechazados sin revisión de los referís a discreción del editor. Si el mismo es conveniente, el manuscrito se enviará para su revisión por pares.
- El editor tomará la decisión de aceptar el artículo basada en los comentarios recibidos por los referís.
- La decisión del editor es definitiva e inapelable.

La figura 1 muestra el proceso de revisión de pares evaluadores empleado por la revista Killkana Técnica.

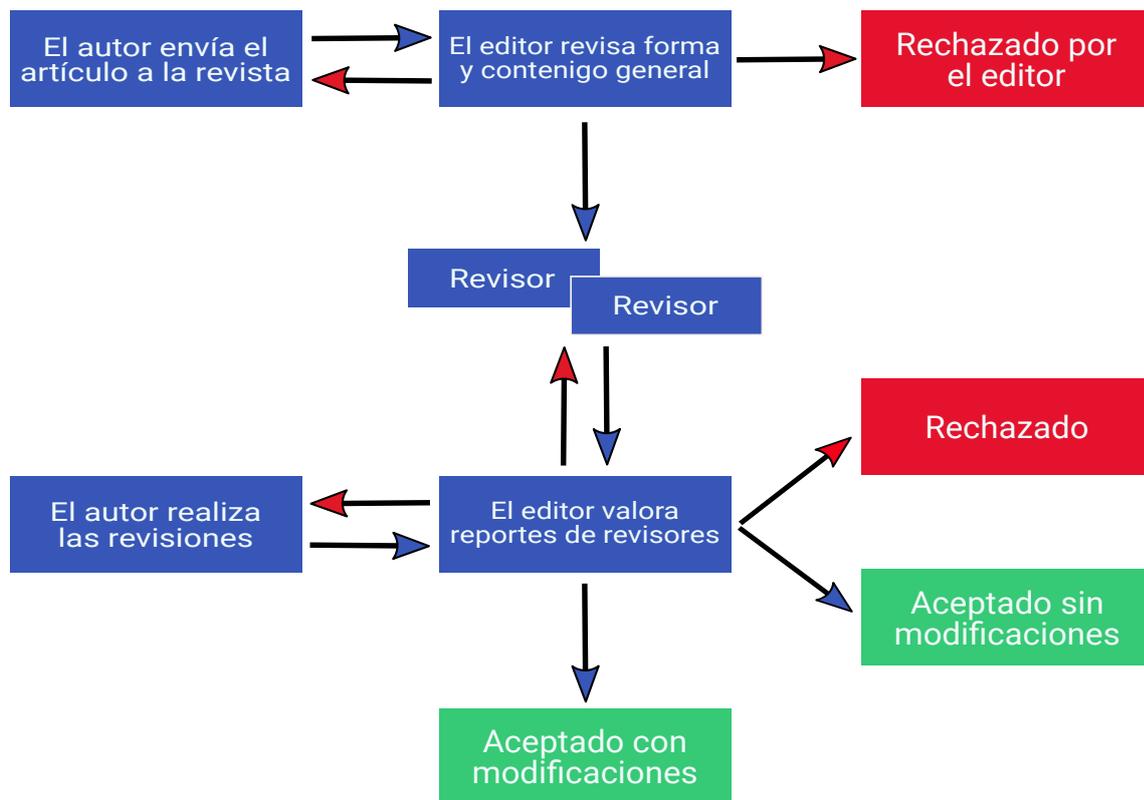


FIG. 1. Diagrama del proceso de revisión de pares evaluadores de las revistas Killkana Técnica.

Hay que aclarar, que luego de que el artículo ha sido aceptado, será sometido a una revisión de estilo, donde se darán los toques finales respecto a la redacción, resumen en idioma inglés, etc.

III. QUIENES SE CONSIDERAN PARA SER PARES EVALUADORES

Los pares evaluadores son investigadores académicos y profesionales, que trabajan en el campo de conocimientos del artículo. Están familiarizados con la literatura científica y tienen contribuciones y/o artículos en el área. Tales expertos no pertenecen a la Universidad Católica de Cuenca.

Los referís ofrecen su tiempo y experiencia de forma voluntaria para mejorar la calidad de los artículos de las revistas y alentar nuevas investigaciones en sus respectivas áreas del conocimiento.

IV. OBJETIVOS DE LOS PARES EVALUADORES

El editor conjuntamente con los pares evaluadores deben asegurar en la medida de lo posible que los trabajos tengan las siguientes características:

- El artículo es un trabajo original, el cual no ha sido publicado previamente ni está en consideración de otra revista, total o parcialmente.

- El manuscrito tiene los estándares de ética.
- El artículo es relevante para los propósitos, alcance y lectores de la revista.
- El trabajo representa resultados originales.
- El manuscrito de revisión enviado ofrece un estudio amplio, completo y crítico, y evalúa la fuentes claves de literatura para un tópico específico.
- El artículo es metodológica y técnicamente bueno.

V. TIEMPOS EN EL PROCESO DE ACEPTACIÓN

La revisión por pares es un proceso riguroso, que demanda su debido tiempo y atención. Entre los factores que pueden influir en el tiempo de respuesta de un trabajo se pueden considerar:

- Los manuscritos son revisados en primera instancia por algunos miembros del equipo editorial.
- En la fase de la revisión por pares, la demora ocurre inevitablemente cuando los referís están ocupados. Como expertos en el área y miembros de instituciones de educación e investigación superior, su tiempo es compartido con otras actividades.
- Las revistas pueden tener un gran número de sumisiones de trabajos.

- El equipo editorial de las revistas Killkana trabaja para asegurar que el proceso de revisión por pares sea riguroso y a tiempo. Los tiempos de espera pueden variar dependiendo de la revista, el campo de investigación, etc.

