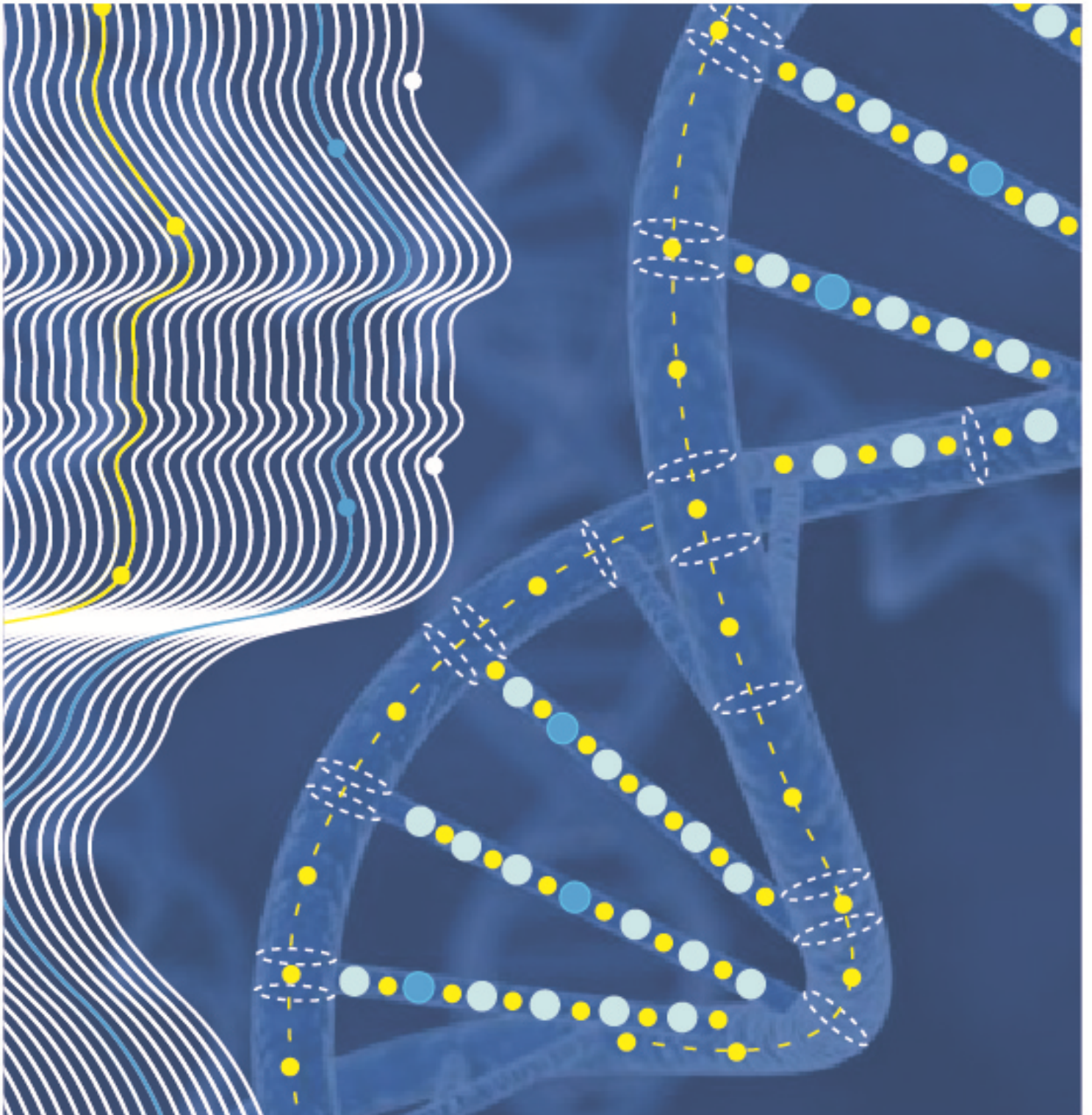




Pontificia Universidad
Católica del Ecuador

Ciencia, Tecnología y Sociedad

ISBN: 978-9942-9902-9-7



Copyright: Dirección de Investigación de la Pontificia Universidad Católica del Ecuador Sede Esmeraldas, prohibida la reproducción total o parcial de este libro por ningún medio impreso o electrónico sin el permiso previo y por escrito del dueño del copyright.

PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR SEDE ESMERALDAS

Dirección: Espejo y subida a Santa Cruz

Casilla: 08-01-0065

Teléfonos: +593 (06) 2721983 – 2721595

Email: direccion.sistemas@pucese.edu.ec

www.pucese.edu.ec

ESMERALDAS - ECUADOR

PUBLICACIÓN ELECTRÓNICA

ISBN:

ISBN: 978-9942-9902-9-7



Los artículos incluidos en esta publicación fueron sometidos a procedimientos de admisión y revisión por pares, llevados a cabo por un comité científico nacional e internacional de alto nivel.

SEGUNDAS JORNADAS DE INVESTIGACIÓN. CIENCIA, TECNOLOGÍA Y SOCIEDAD

Coordinador de la Carrera de Sistemas y Computación

Víctor Xavier Quiñonez Ku

Director de Investigación PUCE Esmeraldas

Ignacio Carazo Ortega

Comité Organizador

Víctor Xavier Quiñonez Ku

Jaime Paúl Sayago Heredia

Wilson Gustavo Chango Sailema

Marc Grob

Manuel Rogelio Nevarez Toledo

Comité Editorial

Pablo Antonio Pico Valencia, *Universidad de Granada – España*

Cesar Raúl García Jacas, *Universidad Nacional Autónoma de México*

José Luis Sampietro Saquicela, *Universidad de Barcelona – España*

Juan Luis Casierra Cavada, *Pontificia Universidad Católica del Ecuador*

Pedro Roberto Suarez Suri, *Pontificia Universidad Católica del Ecuador*

Víctor Xavier Quiñonez Ku, *Pontificia Universidad Católica del Ecuador*

Jaime Paúl Sayago Heredia, *Pontificia Universidad Católica del Ecuador*

Luis Alberto Herrera Izquierdo, *Pontificia Universidad Católica del Ecuador*

Franyelit María Suarez Carreño, *Pontificia Universidad Católica del Ecuador*

Evelin Lorena Flores García, *Pontificia Universidad Católica del Ecuador*

Cristobal Baixauli Pérez, *Pontificia Universidad Católica del Ecuador*

Comité Técnico

Susana Gabriela Patiño Rosado

José Luis Carvajal Carvajal

Pedro César Godoy Rosero

Junior Williams Mera Quiroz

Kleber Rolando Posligua Flores

Víctor Fabián Martínez Estupiñán

Walter Eduardo Almeida Clavijo

Verónica del Carmen Yáñez Ortiz

Ángel Anchundia Ortiz

Rhay Pablo Bustos Vera

Zully Betty Carvache Franco

Luis Enrique Hidalgo Solorzano

Walter Clemente Mosquera Torres

Tania Elizabeth Posligua Flores

Kleber Leonardo Vera Tortorella

Diseño Gráfico

Melba Cristina Marmolejo Cueva

Autor: PUCE Sede Esmeraldas

Editorial: Pontificia Universidad Católica del Ecuador Sede Esmeraldas

Colección: Investigación, Temas relacionados con la Tecnología

Fecha publicación: Enero 2019

Edición: 2da edición

Idioma: Español

PRESENTACIÓN

II Jornada de investigación, ciencia, tecnología y sociedad



Quiero felicitar a los docentes de la carrera de Sistemas y la nueva carrera de TIC por organizar estas segundas jornadas de investigación, ciencia, tecnología y sociedad.

Somos conscientes de la importancia de socializar y compartir experiencias e investigaciones en este campo de las tecnologías que atraviesa toda actividad humana. Más cuando permanentemente se están dando avances sorprendentes.

Para la mayoría, los avances tecnológicos que disponemos actualmente eran impensables hace sólo 30 años. Con el cambio acelerado en el que estamos inmersos ¿Cómo será la tecnología en el 2015? Seguramente la tendremos incorporada en nuestro propio cuerpo.

Por supuesto que tendrá que ser una tecnología sostenible. Ya hace tiempo que sabemos que los recursos no son ilimitados y muchas sociedades viven a “crédito” de recursos que no se sabe si podrán pagar.

Quiero subrayar la importancia de que el título de las jornadas contenga la palabra sociedad. La sociedad nos refiere a las personas y a la ética. No hay que perder de vista que la tecnología no tiene entidad en sí misma. Por ello deberá ser una tecnología sin brechas, que esté a disposición de todos y todas, que respete la dignidad de las personas, y que nos haga la vida más fácil y más feliz.

Un cordial saludo

Aitor Urbina

PRO-RECTOR

ÍNDICE

STACKS FOR WEB APPLICATIONS BASED ON REST SERVICES: MEAN VS JAVA EE	6
ROBOT ASISTENTE PARA CONDUCCIÓN VEHICULAR	14
EDUCCIÓN DE PROCESOS DE NEGOCIO BASADA EN GENERACIÓN AUTOMÁTICA DE PREGUNTAS	23
IMPACTO EN LA AGILIDAD DE LA ANALÍTICA DE DATOS USANDO TECNOLOGÍA IN-MEMORY	33
REVISIÓN SISTEMÁTICA DE LITERATURA: ANÁLISIS DE RIESGOS UTILIZANDO REDES BAYESIANAS	41
EL RETO DE LA TECNOLOGÍA DE LA INFORMACIÓN Y COMUNICACIÓN PARA LA EDUCACIÓN EN EL ECUADOR	50
SISTEMA SCADA PARA MONITOREO DE CONDICIONES FÍSICO Y QUÍMICAS DE UN ACUARIO DE PECES TROPICALES	55
MÉTODO DE SEGURIDAD EN LA AUTENTIFICACIÓN DE CONTRASEÑA UTILIZANDO AES	59
ANÁLISIS COMPARATIVO DE METODOLOGÍAS ÁGILES DE DESARROLLO DE SOFTWARE: UNA REVISIÓN BIBLIOGRÁFICA	65

Stacks para aplicaciones web basados en servicios REST: MEAN vs Java EE

Stacks for web applications based on REST services: MEAN vs Java EE.

Jaime Sayago Heredia, Gustavo Chango Sailema, Evelin Flores, Patricia Medina Peñaloza

Carrera de Sistemas y Computación, Pontificia Universidad Católica del Ecuador. Sede Esmeraldas, Sede Esmeraldas, jaime.sayago@pucese.edu.ec.

Carrera de Sistemas y Computación, Pontificia Universidad Católica del Ecuador. Sede Esmeraldas, Sede Esmeraldas, wilson.chango@pucese.edu.ec.

Carrera de Sistemas y Computación, Pontificia Universidad Católica del Ecuador. Sede Esmeraldas, Sede Esmeraldas, evelin.flores@pucese.edu.ec.

Analista Técnica del Sistema de Información Local, GAD Intercultural del Cantón Cañar, patricia.medina@canar.gob.ec

Resumen— El uso de los servicios web REST (Representational State Transfer Protocol) ha aumentado y se ha convertido en la implementación más utilizada en la actualidad. La parte más crucial en un proyecto de desarrollo web basado en servicios REST es la elección de las herramientas correctas para el front-end, back-end y entorno de base de datos. El objetivo principal de este trabajo es realizar la comprobación del stack MEAN como solución efectiva para una aplicación web basada en servicios REST. JavaScript ha sido el lenguaje de programación del lado del cliente que se ejecuta en el navegador. La stack MEAN está conformada por MongoDB (base de datos), Node.js. (servidor web), Express (back-end) y Angular (front-end) para construir aplicaciones web. Comparamos la stack MEAN con una stack muy popular como lo es Java EE que agrupa a MongoDB (base de datos), Apache/Tomcat (servidor web), Spring boot (back-end) y JSP/HTML/CSS (front-end), con respecto a sus componentes, características y configuración del entorno. Desarrollamos dos aplicaciones similares construidas con las stacks MEAN y Java EE. Comparamos y analizamos varias características y criterios para pruebas como el tiempo de respuesta, rendimiento y carga de datos

Palabras Clave: *Arquitectura software, Javascript, REST, Servicio Web, Java EE*

Abstract-- The use of REST (Representational State Transfer Protocol) web services has increased and has become the most widely used implementation today. The most crucial part of a web development project based on REST services is the choice of the right tools for the front-end, back-end and database environment. The main objective of this work is to test the MEAN stack as an effective solution for a web application based on REST services. JavaScript has been the client-side programming language running in the browser. The MEAN stack consists of MongoDB (database), Node.js. (web server), Express (back-end) and Angular (front-end) to build web applications. We compare the MEAN stack with a very popular stack such as Java EE that groups

MongoDB (database), Apache/Tomcat (web server), Spring boot (back-end) and JSP/HTML/CSS (front-end), with respect to its components, features and environment configuration. Two similar applications were built with the MEAN and Java EE stacks. We compared and analyzed several features and test criteria such as response time, performance and data loading.

Keywords: *Software Architecture, Javascript, REST, Web Service, Java EE*

I. INTRODUCCIÓN

La sociedad ha cambiado hacia una mayor dependencia de las tecnologías web que han llegado a convertirse en parte de nuestra vida diaria. Motivando a la academia y a la industria a buscar soluciones para mejorar su disponibilidad, rendimiento y escalabilidad. De esta evolución surgieron los llamados servicios web que se podrían definir como un módulo de software ofrecido por un proveedor de servicios, disponible a través de la web y que es un elemento clave en la integración de sistemas de diferentes plataformas, lenguajes de programación y tecnologías [1]. Una de las implementaciones más utilizadas en la actualidad para servicios web es REST (Protocolo de transferencia de estado representacional) o RESTful. REST fue creado por Roy Fielding y propone un estilo de arquitectura orientada al diseño de software basado en la red [2]. Se puede definir como una arquitectura cliente/servidor. El cliente envía la solicitud al servidor, el servidor procesa la solicitud y devuelve la respuesta. Estas solicitudes y respuestas se basan en la transferencia de representaciones de recursos que se identifica a través de URIs [3]. El uso de servicios web REST han ido en aumento, dando surgimiento a varias herramientas, frameworks y stacks para desarrollo, integración y ejecución de aplicaciones web, las más utilizadas en la actualidad son la stack MEAN con Javascript y la stack de Java EE. El problema radica que al

momento de elegir la arquitectura apropiada de tecnologías para desarrollo, integración y ejecución de aplicaciones web basados en tecnologías conocidas y probadas como son los servicios web REST. La elección del stack correcto no es fácil, ya que muchos parámetros tienen que considerarse en la base de datos, en el lado del servidor o el lado del cliente, etc. Diversos autores presentan a la stack MEAN como la mejor solución [4] [5] [6]. En este trabajo se plantea una comparación entre la stack MEAN y la stack Java EE y proponer cuál stack es la solución más efectiva para el desarrollo de aplicaciones web basados en servicios web REST. Se realizó la implementación de una aplicación para el manejo de la hoja de vida y gestión docente para el departamento de talento humano de la Pontificia Universidad Católica del Ecuador PUCE, Sede Esmeraldas – Ecuador, que permite almacenar, gestionar y analizar la información de los docentes y su gestión docente semestral. Esta aplicación web fue construida utilizando los dos stacks para probar y analizar los pros y contras de los dos stacks, junto con algunas pruebas de evaluación comparativa.

II. ARQUITECTURAS DE SOFTWARE

Cada sistema informático, grande o pequeño, está formado por piezas que están unidas entre sí. Puede haber un pequeño número de estas piezas, o tal vez solo una, o puede haber docenas o cientos; y este vínculo puede ser trivial, o muy complicado, o en algún punto intermedio, es decir cada sistema tiene una arquitectura [7]. Al tratar de definir el concepto de Arquitectura de Software existen varias definiciones alternativas o contrapuestas. Pero hay algunas que son reconocidas, a continuación, las citamos. Se puede definir la arquitectura de software como la descomposición de un sistema de nivel superior en cada uno de sus componentes principales, las interfaces y su comunicación [9]. La IEEE la define de la siguiente manera: Es la organización fundamental de un sistema encarnada en sus componentes, las relaciones entre ellos y el ambiente y los principios que orientan su diseño y evolución [10]. Dentro de las ventajas de la arquitectura de software, podemos mencionar que: simplifica la capacidad de comprender sistemas complejos planteando limitaciones en el diseño, reutiliza componentes en múltiples niveles, separa elementos, lo que permite mayor facilidad en el cambio y mantenimiento del desarrollo, aplica utilización de patrones [11]. El uso de arquitectura de software puede realizarse de una manera eficiente en función del tiempo y ser rentable ya que permite la reutilización de los componentes y patrones de diseño en los proyectos [12]. Al implementar una arquitectura de software es fundamental el uso de los patrones arquitectónicos que brindan un principio general de estructura. Un patrón arquitectónico expresa un esquema de organización estructural fundamental para los sistemas de software. Proporciona un conjunto de subsistemas predefinidos, especifica sus responsabilidades e incluye reglas y pautas para organizar las relaciones entre ellos [13].

A. REST

El estilo de Arquitectura REST utiliza identificador de recursos uniforme URI para acceder y manipular los estados de los recursos en cual se proporciona un identificador único para un

recurso web [3]. REST normalmente tiene muchos recursos y es una arquitectura orientada a los recursos [14]. En REST se acceden a los recursos mediante una interfaz genérica la cual reduce dramáticamente la complejidad de la semántica de la interfaz durante la interacción de servicios. Las necesidades de ancho de banda y recursos de REST son bajos, por lo que se considera ligero [15]. Las aplicaciones desarrolladas utilizando REST son llamadas aplicaciones RESTful que proporcionan más flexibilidad y menos gastos generales, lo que resulta en una mejor solución para la mayoría de las implementaciones [16].

B. SERVICIOS WEB

El término servicios web según la W3C (World Wide Web Consortium) proporcionan un medio estándar de interoperabilidad entre diferentes aplicaciones de software que se ejecutan en una variedad de plataformas y/o marcos de trabajo [17]. Un servicio web es un sistema de software diseñado para soportar la interacción entre máquinas a través de una red. Tiene una interfaz descrita en un formato procesable por máquina. Otros sistemas interactúan con el servicio web de una manera prescrita por su descripción utilizando mensajes, típicamente transmitidos usando HTTP con una serialización en conjunción con otros estándares relacionados con la web [17]. Se podría entender como un conjunto de aplicaciones o tecnologías que pueden interoperar en la web, su principal función es la de intercambiar datos entre sí; tanto los proveedores ofrecen sus servicios y los usuarios solicitan el servicio llamando a estos procedimientos utilizando la web [18]. Estos servicios proveen de mecanismos de comunicación estandarizados entre las distintas aplicaciones, que pueden interactuar entre ellas para entregar la información al usuario. La interoperabilidad y extensibilidad de las aplicaciones y su combinación para operaciones complejas requiere la utilización de un estándar [17].

C. ARQUITECTURA REST

REpresentational State Transfer (REST) es un estilo arquitectónico propuesto por Fielding [19]. REST es para sistemas hipermedias distribuidos a gran escala y que logra que la World Wide Web (WWW) sea escalable. Fielding argumenta que en REST es la existencia de recursos (elementos de información), que pueden ser accedidos utilizando un identificador global (un Identificador Uniforme de Recurso). Para manipular estos recursos, los componentes de la red (clientes y servidores) se comunican a través de una interfaz estándar (HTTP) e intercambian representaciones de estos recursos (los ficheros que se descargan y se envían). El cliente puede navegar esencialmente a través de una amplia gama de recursos existentes, siguiendo los enlaces de un recurso a recurso [20]. Un principio clave de REST es la interacción sin estado entre los participantes en una conversación. Un estado en este caso significa el estado de la aplicación/sesión. El estado se mantiene como parte del contenido transferido del cliente al servidor/servicio y viceversa [21]. Más concretamente, en el caso de los servicios, los clientes que desean utilizar un servicio acceden a una representación particular de los recursos que representan el servicio mediante la transferencia de contenido de la aplicación utilizando un

conjunto pequeño y definido globalmente de métodos remotos [21]. Estos métodos describen la acción a realizar sobre los recursos. Los métodos HTTP para crear, leer, actualizar y borrar recursos, cada uno identificado por un URI, son (PUT, POST, GET, y DELETE en HTTP 1.0, mientras que HTTP 1.1 permite extensiones) [22]. En REST, cada solicitud enviada a un objeto resulta en la transferencia de una representación de este objeto por ejemplo, texto, XML, JSON, etc. [20]. REST se ha convertido en la implementación más utilizada en la actualidad [23].

III. STACK MEAN

La stack MEAN es una solución potente y completa. Utiliza el lenguaje de programación JavaScript [24]. Comprende cuatro bloques principales: MongoDB como base de datos, Express como marco de trabajo del servidor web, AngularJS como marco de trabajo del cliente web y Node.js como plataforma del servidor. Estos componentes están desarrollados por diferentes equipos e involucran a una comunidad fuerte de desarrolladores y defensores impulsando el desarrollo y documentación de cada componente. Sin embargo, un problema que podría afectar dramáticamente su proceso de desarrollo y presentar problemas de escalamiento y arquitectura es la conexión de estas herramientas [25]. La principal fortaleza de la pila MEAN radica en su centralización de JavaScript como el principal lenguaje de programación, ya que cada componente está escrito en JavaScript, incluso la base de datos almacena los datos en formato JavaScript Object Notation (JSON) que es el único script que JavaScript entiende completamente [26]. Por lo tanto, JavaScript no sólo se utiliza como lenguaje de scripting del lado del cliente, sino que también se utiliza a lo largo de la aplicación, es decir, en el lado del cliente, del servidor y de la base de datos [4]. El uso de JavaScript como lenguaje principal de programación tanto en el lado del cliente como en el lado del servidor hace que la pila MEAN sea más potente y reduce el tiempo en la construcción de la aplicación [5]. El uso de todo el JavaScript permite dividir la funcionalidad y las tecnologías utilizadas de la siguiente manera [26]:

- Base de datos objetos JSON utiliza MongoDB.
- Servidor web utiliza Node.js.
- Framework web (back-end) utiliza Express.
- Cliente (front-end) utiliza Angular.

A. MONGODB

MongoDB es un modelo de almacenamiento de documentos NoSQL potente, adaptable y escalable [27]. En MongoDB, los datos se almacenan en la base de datos en un formato JSON llamado BSON, que significa JSON binario, en lugar de filas y columnas como en la base de datos relacional. MongoDB tiene una capacidad para escalar con varias características que la base de datos relacional proporciona como índices secundarios, clasificación y consultas [28], que proporciona alto rendimiento, alta disponibilidad y escalabilidad [25].

B. EXPRESS

Express es un framework web maduro y flexible para construir aplicaciones web sobre el ecosistema Node. Por defecto, el framework Express utiliza el motor Pug para soportar plantillas [29]. Express es un framework relativamente pequeño que se encuentra en la parte superior de la funcionalidad del servidor web de Node para simplificar sus APIs y añadir nuevas funciones útiles. Facilita la organización de la funcionalidad de su aplicación con middleware y enrutamiento; agrega utilidades útiles a los objetos HTTP de Node y el renderizado de vistas HTML dinámicas; define un estándar de extensibilidad fácilmente implementado [30]. Express es fácil de configurar, implementar, controlar y proporcionar varios componentes clave para manejar las solicitudes web. Express ayuda en la creación de aplicaciones web y servidores HTTP simples ya que es un framework mínimo y flexible [30]. En MEAN, Express funciona como un medio para transferir las solicitudes de un cliente a una base de datos y envía las respuestas de la base de datos al cliente [26].

C. ANGULARJS

AngularJS es una librería escrita en JavaScript para el desarrollo de aplicaciones web, mantenida por Google, es un framework JavaScript de código abierto y aborda los retos de las single-page applications (SPAs) [31]. Una aplicación web AngularJS sigue el patrón de diseño MVC, que resulta en el desarrollo de aplicaciones web ampliables, mantenibles, comprobables y estandarizadas [31]. La unión de datos AngularJS y la inyección de dependencias lo convierten en un socio ideal para cualquier tecnología de servidor, ya que elimina gran parte del código que de otro modo tendrías que escribir, y todo sucede dentro del navegador [26].

D. NODE

Node es un framework de desarrollo desarrollado originalmente en 2009 por Ryan Dahl, basado en el motor JavaScript V8 de Google [32]. Node es una plataforma basada en el tiempo de ejecución JavaScript de Chrome para crear aplicaciones en red rápidas y escalables [25]. Node utiliza un modelo controlado por eventos y de no bloqueo de E/S que lo hace ligero y eficiente, perfecto para aplicaciones que requieren grandes cantidades de datos y que se ejecutan en dispositivos distribuidos [26]. Node es un lenguaje de scripting del lado del servidor que puede ser usado en el lado del servidor, lado del cliente, e incluso puede ser un servidor web. Antes de que existiera el Node, JavaScript se usaba simplemente para la interacción del usuario como script del lado del cliente [33].

IV. STACK JAVA EE

Java que es uno de los lenguajes de programación más utilizados y que tiene una variedad de herramientas para diferentes propósitos [34]. En este caso se combina varias herramientas y componentes la conformar la stack Java EE. En una aplicación web los componentes están en el front-end o en el back-end [31]. La stack Java EE que es muy utilizada en para el desarrollo de aplicaciones modernas [35]. Estas tecnologías deben trabajar de una manera fácil y armoniosa. En el Front-End se usa JSP/HTML/CSS, utilizando el patrón

MVC. En el back-end Spring Boot y la base de datos Postgresql, MongoDB o cualquier otro motor de base de datos. El uso de toda la stack Java EE permite dividir la funcionalidad y las tecnologías utilizadas de la siguiente manera, que conforman la stack seleccionada:

- Base de datos MongoDB.
- Spring boot para el (back-end).
- Cliente (front-end) utiliza JSP/HTML/CSS.
- La plataforma Apache/Tomcat como servidor web.
- A continuación, se procede a describir cada una de las tecnologías utilizadas.

A. SPRING BOOT

Spring Boot es un framework que posee las bibliotecas necesarias para la creación de aplicaciones ejecutables basadas en Spring. Los componentes que posee Spring Boot permiten simplificar los pasos de la selección de las dependencias necesarias para el proyecto a desarrollar y su despliegue en el servidor. No requiere de configuración usando archivos en XML/JSON y permite la creación de servicios REST en Java de manera fácil, centrando al desarrollador en los elementos críticos de desarrollo específico de su aplicación [36].

B. SPRING DATA REST

Spring MVC ofrece una base sólida para construir los tipos de servicios web REST que son el medio más utilizado para la integración de aplicaciones en la web, pero la implementación de la funcionalidad básica del servicio web REST puede ser tediosa y dar como resultado una gran cantidad de código [37]. Spring Data REST se basa en los repositorios de Spring Data y los exporta automáticamente como recursos de REST. Aprovecha los hipermedios para permitir a los clientes encontrar la funcionalidad expuesta por los repositorios y permite integrar los recursos en la funcionalidad relacionada basada en hipermedios tan fácilmente como sea posible. Spring Data REST es en sí misma una aplicación Spring MVC y está diseñada de tal manera que debe integrarse con sus aplicaciones Spring MVC existentes con muy poco esfuerzo. Una capa de servicios existente (o futura) puede funcionar junto con Spring Data REST con cambios menores.

C. GRADLE

Es una herramienta de compilación basada en JVM, usa un lenguaje potente y expresivo Domain Specific Language (DSL) implementado en Groovy en lugar de XML. Debido a que Gradle es nativo de JVM, permite escribir una lógica personalizada sea en el lenguaje Java o Groovy. Sin embargo, escribir los archivos de construcción de Gradle no requieren un conocimiento profundo de Groovy [38]. Gradle tiene mayores ventajas a diferencia de Maven o Ant, una de ellas es que Gradle no usa XML, en cambio Maven y ANT si lo usan, lo cual queda algo corto en expresar el flujo del programa y la lógica condicional. Gradle brinda una sofisticada gestión de dependencias robusta y potente y son muy adecuados para construir proyectos basados en Java, debido a que permite sin mayor esfuerzo el consumo de componentes o módulos desarrollados por terceros o definir su propia biblioteca de dependencias por proyecto permitiendo construir aplicaciones

escalables. Brinda muchos beneficios, el más importante es su flexibilidad. [39].

D. JSP/HTML/CSS

HTML (HyperText Markup Language) que en español significa Lenguaje de Marcas de Hipertexto es considerado un lenguaje de marcado principal de la World Wide Web que originalmente se diseñó para detallar documentos científicos semánticamente, sin embargo, al paso de los años se ha podido adaptar para detallar y describir otros tipos de documentos como por ejemplo aplicaciones [40]. HTML se encuentra en su versión 5.2 que provee diferentes características como estructura, funcionalidad, adaptabilidad y estilo; al hablar de este último, CSS (Cascading Stylesheets) se destaca que estas tecnologías son altamente dependiente entre sí [41]. Java Server Page o mejor conocida como JSP es una página HTML que brinda la ventaja de poder incrustar código Java facilitando así la creación de páginas web dinámicas [42]. Una de las principales ventajas que ofrece JSP ante otros lenguajes es que con él se pueden crear aplicaciones web que manejen una lógica de negocios y acceso a datos de manera fácil y sencilla, de tal manera que permita la separación de los niveles dentro de la aplicación. Otra ventaja es la posibilidad de heredar la portabilidad del lenguaje Java [42].

V. STACK MEAN Y JAVA EE: CARACTERÍSTICAS, VENTAJAS Y DESVENTAJAS

En este apéndice se muestran las características, ventajas y desventajas de cada una de las stacks. MEAN gana popularidad debido a las características de sus componentes: Node.js, MongoDB, AngularJS, junto con Express y todo utilizando JavaScript que es una de sus mayores ventajas. Aunque podemos afirmar que MEAN no se recomienda en sistemas pesados de procesamiento computacional, debido a que el componente Node.js es un entorno de un solo hilo y esto sería un inconveniente fuerte de esta tecnología. La madurez de la tecnología Java EE es la opción más correcta para este tipo de aplicaciones pesadas de procesamiento computacional. Pero para una aplicación de una página (SPA), que requiere una alta interacción con el usuario, visualización y alta escalabilidad, la pila MEAN es la mejor opción. Cabe recalcar que en el ámbito empresarial Java EE es la tecnología más popular para el desarrollo de aplicaciones web por la seguridad que brinda esta tecnología. La elección de la stack depende principalmente de algunas características, que las vamos a analizar para conformar un conjunto de parámetros que permiten realizar la comparativa entre estas dos stacks, en definitiva, se trata de buscar un marco de evaluación de la arquitectura con los criterios elegidos. A continuación, se detalla los parámetros elegido y al final encontramos un resumen de las stacks a partir de la elección de parámetros efectuada. La siguiente tabla muestra un resumen de las características de cada stack. Podemos decir que ambas stacks tanto MEAN y Java EE, cumplen con las características evaluadas manejándolas de distinta forma a través de distintas tecnologías. La stack MEAN resuelve de mejor manera la mayoría de las características evaluadas antes que la stack Java EE.

Tabla 1.
Resumen de características de las stacks.

Característica	MEAN	Java EE
Capa de servidor simplificada	Construcción de servidor localmente	Empaquetamiento de artefacto en un servidor embebido
Código isomórfico	Código puede ser ejecutado en el front-end como el back-end	Código puede ser ejecutado en el front-end como el back-end
Escalabilidad	La aplicación puede expandirse los eventos se ejecutan libremente de manera asíncrona	La aplicación se expande a través del servidor embebido con una distribución multihilos.
Arquitectura sin bloqueo	Maneja el uso de bucles de eventos sin bloqueo en un solo hilo.	El bloqueo se resuelve a través de múltiples hilos de ejecución
Tiempo de desarrollo	Al utilizar el mismo lenguaje JavaScript se reduce el esfuerzo y tiempo de desarrollo	Al utilizar el mismo lenguaje Java se reduce el esfuerzo y tiempo de desarrollo. Existen varios marcos de trabajo para mejorar el tiempo de desarrollo
Transformación de datos y extensibilidad	Los datos manejados para el intercambio de información es JSON.	Los datos manejados para el intercambio de información por estándar son XML, pero se puede trabajar con JSON.

VI. CASO DE ESTUDIO: COMPARACIONES, PRUEBAS Y ANÁLISIS

En esta sección, se presentan las pruebas junto con la interpretación de resultados de comparativas y pruebas de la aplicación web de hoja de vida y gestión docente a través de servicios REST. Para probar y analizar los pros y contras de las dos stacks, se desarrolló la aplicación de la hoja de vida y gestión de docente para el departamento de talento humano de la Pontificia Universidad Católica del Ecuador Sede Esmeraldas, que fue construido utilizando ambas pilas, seguido de algunas pruebas de evaluación comparativa. Algunos de los criterios como el desarrollo de aplicaciones en tiempo de respuesta, el rendimiento, la carga de datos son demostrados en las siguientes secciones. Las pruebas de la aplicación fue con la ayuda de tres herramientas de evaluación comparativa: Siege [43], Apache Bench [44] y de tiempo de respuesta fue JMeter [45]. La aplicación que se desarrolló es un CRUD, con varios datos del docente para llenar su hoja de vida por ejemplo datos personales, títulos, experiencia docente, experiencia no docente, capacitaciones, investigaciones, artículos, congresos, libros, etc. Junto con un informe de actividades de fin de cada semestre con la información de alumnos tutorados, cursos, seminarios, congresos y proyectos de vinculación realizados. El desarrollo de esta aplicación fue de apreciar que tan fácil o complicado puede ser realizar una aplicación utilizando ambas stacks. El objetivo principal era para ilustrar la el tiempo de respuesta tanto en la stack MEAN como en la stack Java EE, como también medir su rendimiento y carga de datos. Para el propósito de la prueba, alrededor de 300 registros de docentes y de gestión fueron insertados en las dos aplicaciones.

A. PRUEBA DE TIEMPO RESPUESTA DE LAS APLICACIONES

Las mediciones de tiempo de respuesta de las aplicaciones construidas con los dos stacks MEAN y Java EE, se han realizado mediante Apache JMeter. El tiempo de respuesta significa el intervalo entre el momento en que se envía una solicitud y el momento en que JMeter recibe la respuesta [46]. Para esta prueba se desarrolló un script en JMeter para la creación de la petición de datos. Básicamente, el script intenta acceder a la aplicación y realizar una función de búsqueda de información. El experimento se realizó en el entorno donde glassfish esto en la stack Java EE y para la stack MEAN el entorno Node.js. El script de JMeter se lo ejecuto por separado. La siguiente tabla muestran los resultados de las dos pruebas.

Tabla 2.

Tiempo de respuesta de las dos stacks MEAN y Java EE.

Etiqueta	Mues tras	Prome dio	Error %	Rendi miento	Kb/seg Recibidos
Java EE	1000	953	0,00%	350,30%	2351,28
MEAN	1000	2	0,00%	352,50%	412,12

Como se aprecia en la tabla se hicieron 1000 solicitudes de muestras. El tiempo promedio transcurrido (columna Promedio) fue de 953 milisegundos y el porcentaje de error (columna Error %) fue de 0 para la stack Java EE. Para la stack MEAN el tiempo promedio transcurrido (columna Promedio) fue de 2 milisegundos y el porcentaje de error (columna Error %) fue de 0. El porcentaje de error es la columna que muestra el porcentaje de solicitudes rechazadas de esta tabla. En ambas stacks esta columna es la que no varía entre ejecuciones. El resto de las columnas muestran la rapidez con la que se manejaron las solicitudes (columna de rendimiento). Respecto al rendimiento hay una pequeña diferencia favorable para la stack Java EE 350.30% frente a 352.50% para la stack MEAN. Por último, la cantidad de kilobytes recibidos (columna de

kb/segundo recibidos). En esta podemos apreciar que la stack Java EE su velocidad es de 2351.28 kb/segundo frente a 412.12 kb/segundo de la stack MEAN. A continuación, la gráfica de la prueba realizada con el script de JMeter.

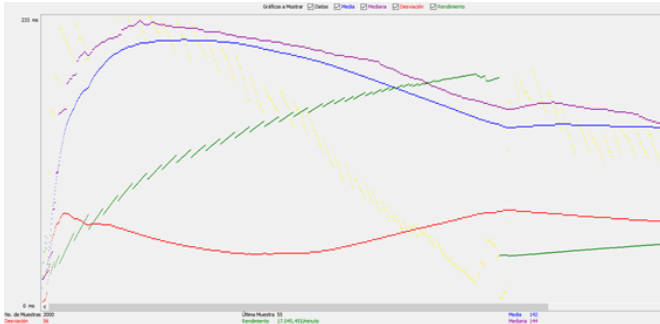


Fig. 1. Gráfica JMETER prueba tiempo de respuesta.

B. PRUEBAS DE RENDIMIENTO

Para esta prueba, usamos la herramienta Apache Bench para probar el tiempo de respuesta del sistema. La prueba se ejecutó con varios niveles de concurrencia y número de peticiones. La concurrencia es la medida de cuántas sesiones de usuario simultáneas están activas en una aplicación web en un momento dado. Con un aumento en la concurrencia, podemos probar cuántas sesiones de usuarios concurrentes puede soportar una aplicación web en términos de tiempo de respuesta para realizar cualquier tarea. Este es un comando de ejemplo para comparar un servidor que se ejecuta localmente en el puerto 8080 con 10.000 solicitudes en total y 500 solicitudes concurrentes: “ab -n 10000 -c 500 [http://localhost:8083/ GestionDocenteProyecto/](http://localhost:8083/GestionDocenteProyecto/)”. Se realizaron varias pruebas con aumentos en la concurrencia y el número de solicitudes como se muestra en la Tabla 3. Las pruebas se realizaron para MEAN y Java EE con la concurrencia establecida en 10, 50, 100 y 500, y el número de solicitud como 100, 500, 1000, 5000, 10000, 50000 y 100000.

Tabla. 3.

Tiempo de respuesta en milisegundos de 100000 solicitudes.

Concurrencia	10	50	100	500
JAVA (respuesta en ms)	64.216	180.072	224.170	1.451.031
MEAN (respuesta en ms)	8.395	34.006	68.301	697.426

Observamos en la tabla que es la prueba de 100000 respuestas y un nivel de concurrencia de 10, 50, 100 y 500. En la que para la stack Java EE el tiempo medio que el servidor ha tardado en atender un grupo de peticiones fue de 64.216, 180.072, 224.170 y 1.451.031 milisegundos respectivamente. Y para MEAN este tiempo de fue de 8.395, 34.006, 68.301 y 6097.426 milisegundos. Se puede apreciar que mientras aumenta el número de solicitud de respuesta y concurrencia mejora el rendimiento de la stack MEAN sobre el stack Java EE.

C. PRUEBAS DE TRANSFERENCIA DE DATOS Y CONCURRENCIA

Prueba de transferencia de datos y concurrencia utilizando Siege. Para probar cargas en ambas aplicaciones construidas en las stacks MEAN y Java EE, la función Siege se utilizó un instrumento de evaluación comparativa. A continuación, se muestra un comando de ejemplo para comparar un servidor que se ejecuta localmente en el puerto 4200 con 150 peticiones simultáneas, durante un lapso de 5 minutos [47] para la stack MEAN:

```
siege --concurrent=150 --time=5M http://localhost:4200.
```

Para la prueba la realización de la prueba para la stack Java EE, se utiliza la ejecución localmente hacia el puerto 8083 donde está alojada la aplicación web. La prueba realizada por Siege proporciona el tiempo transcurrido, la cantidad de datos transferidos, tiempo de respuesta, rendimiento y tasa de transacciones. Estos datos fueron recolectados de Siege pasando el valor de concurrencia y el URL de la aplicación como entrada. El resultado de la prueba de asedio para los datos transferidos para las stacks Java EE y MEAN se muestra en las siguientes tablas y junto a sus respectivos análisis. En la siguiente tabla se resume la tasa de transacciones promedio de transacciones que el servidor fue capaz de manejar por segundo. Con los distintos niveles de concurrencia de 10, 20, 50, 100 y 500. Esta tasa fue de es menor de transferencias por segundo para MEAN y es superior los datos de transferencia por segundo para Java EE.

Tabla 4.

Datos transferidos durante la concurrencia.

Concurrencia (MB)	5	10	20	50	100	150
MEAN	2,66	5,3	10,61	26,57	53,26	79,36
Java EE	18,48	34,29	74,25	185,09	365,46	547,32

Tabla 5.

Datos de concurrencia de las stack MEAN y Java EE.

Concurrencia	5	10	20	50	100	150
MEAN	0,03	0,06	0,13	0,41	0,72	1,04
Java EE	0,08	0,16	0,34	0,93	2,17	3,97

Como lo podemos observar en la tabla 5, se observa de mejor manera la prueba de concurrencia que es mayor para la stack Java EE. Mientras que la stack MEAN es menor. Esta medida es un parámetro importante en el estudio. Por lo que podemos afirmar que el servidor tiene un mejor rendimiento en la stack MEAN que el servidor de la stack Java EE.

VII. RESULTADOS, CONCLUSIONES, RECOMENDACIONES Y LIMITACIONES

En el estado del arte se hizo una descripción de las tecnologías para arquitecturas y servicios web, describió y analizó dos tipos de tecnologías de stacks para el desarrollo de web aplicaciones utilizando servicios REST: la stack MEAN y la stack Java EE. La stack MEAN es un paquete de software que combina MongoDB como la base de datos NoSQL, Express como un framework de Node.js para el scripting para el desarrollo del

back-end, Angular como plataforma MVC para la construcción del front-end, construido con código JavaScript. La stack Java EE es la combinación de MongoDB como el sistema de base de datos, Spring Boot para el back-end y para el front-end se utiliza JSP/HTML/CSS, y Apache/Tomcat como servidor web. En la stack MEAN, JavaScript es el único lenguaje de programación tanto para el front-end y para el back-end. En la stack Java EE, utiliza el lenguaje de programación java del lado del back-end y del lado del front-end junto con HTML/CSS. La creciente popularidad del uso de JavaScript como secuencias de comandos del front-end y del back-end ha hecho que MEAN sea una las combinaciones de tecnologías más utilizadas para desarrollo aplicaciones web y servicios REST. La stack MEAN está construida exclusivamente en JavaScript, por lo que es un lenguaje para gestionar el lado del cliente, servidor y base de datos. Todos los componentes utilizados en la stack MEAN son completamente de código abierto y actualmente son soportados por desarrolladores corporativos como MongoDB y Google. Todos los componentes de MEAN son relativamente ligeros. La stack MEAN es flexible y escalable. Angular es una muy buena opción para el desarrollo de una Simple Page Application y es responsiva. Comparamos la stack MEAN con otra stack muy popular, la stack Java EE, que tiene ya trabajando mucho tiempo con éxito durante décadas. Java EE también es muy fácil de aprender ya que tanto el front-end como el back-end están contruidos usando un solo lenguaje de programación Java. Todos los componentes de Java EE son tecnologías ya maduras. En Java EE, front-end son capaces de entender el código de back-end y los desarrolladores de back-end son capaces de entender el código del front-end, utilizan Java como lenguaje principal y adicionalmente HTML/CSS que son lenguaje de marcado muy populares. JSP es utilizada para el desarrollo de aplicaciones web modernas y es responsiva. Apache/Tomcat es escalable y realiza en el manejo de llamadas asíncronas y en el manejo de muchas tareas concurrentes en un a través de multi-threading. En este estudio, nos centramos en la probar la stack MEAN frente a la stack Java EE con el sistema de la hoja de vida y gestión docente, construida para comparar la stack MEAN contra la stack Java EE. El sistema de la hoja de vida y gestión docente demostró el rendimiento, incluyendo el rendimiento con solicitud de respuesta con concurrencia, transferencia de datos y concurrencia. El rendimiento de la aplicación se midió con la ayuda de herramientas de evaluación comparativa. El tiempo de respuesta de la stack Java EE es mayor que en comparación con la stack MEAN. La transferencia de datos fue más baja en la stack MEAN en comparación con la stack Java EE. Y en la prueba de concurrencia se observa que la concurrencia es mejor en el stack MEAN ya que es menor en comparación con la stack Java EE. En este estudio, nos centramos en la puesta en marcha en una aplicación real con las stacks MEAN y Java EE, dando como resultado que la stack MEAN sobresale en relación con la stack Java EE y en las que, la stack MEAN es preferible a la stack Java EE. Falta hacer mención a situaciones en la que MEAN no debe ser utilizada, puesto que se requiere de mayor cantidad de criterios de análisis y variables de medición. Un análisis de las deficiencias de la stack MEAN nos proporcionaría con una mejor comprensión MEAN y su

aplicabilidad. Sin embargo, este documento demostró cómo la stack MEAN es buena para el desarrollo de aplicaciones web basados en servicios REST, hay limitaciones en nuestro estudio de caso. Por ejemplo, nuestro estudio de caso no muestra cómo MEAN reacciona con las aplicaciones intensivas y de gran procesamiento en CPU y memoria RAM. Por lo tanto, este documento podría beneficiarse de la introducción de un estudio de caso que se ocupe del uso intensivo y gran procesamiento de la CPU y memoria RAM.

REFERENCIAS

- [1] J. Tihomirovs and J. Grabis, "Comparison of SOAP and REST Based Web Services Using Software Evaluation Metrics," *Inf. Technol. Manag. Sci.*, vol. 19, no. 1, pp. 92–97, 2016.
- [2] R. T. Fielding and R. N. Taylor, "Principled design of the modern Web architecture," *ACM Trans. Internet Technol.*, vol. 2, no. 2, pp. 115–150, 2002.
- [3] S. Mumbaikar and P. Padiya, "Web Services Based On SOAP and REST Principles," *Int. J. Sci. Res. Publ.*, vol. 3, no. 5, pp. 3–6, 2013.
- [4] M. Stajcer and D. Orescanin, "Using MEAN stack for development of GUI in real-time big data architecture," *2016 39th Int. Conv. Inf. Commun. Technol. Electron. Microelectron. MIPRO 2016 - Proc.*, pp. 524–529, 2016.
- [5] A. J. Poulter, S. J. Johnston, and S. J. Cox, "Using the MEAN stack to implement a RESTful service for an Internet of Things application," *IEEE World Forum Internet Things, WF-IoT 2015 - Proc.*, pp. 280–285, 2015.
- [6] R. Salunkhe, S. Telang, P. Shrigondekar, and A. Tanpure, *Review of REST Ful Service Using MEAN Stack for Real Time Big Data Architecture*, vol. 3297, no. 11. Birmingham: Packt Publ, 2007.
- [7] N. Rozanski and E. Woods, *Software Systems Architecture: Working with Stakeholders Using Viewpoints and Perspectives*. Addison-Wesley, 2005.
- [8] P. Clements, R. Kazman, and M. Klein, *Evaluating Software Architectures: Methods and Case Studies*. Addison-Wesley Professional, 2001.
- [9] D. Garlan and M. Shaw, "An Introduction to Software Architecture," *Knowl. Creat. Diffus. Util.*, vol. 1, no. January, pp. 1–40, 1994.
- [10] S. Engineering and S. Committee, "IEEE Recommended Practice for Architectural Description of Software-Intensive Systems," 2000.
- [11] D. Garlan and D. Garlan, "Software Architecture : a Roadmap Software Architecture : a Roadmap," 2000.
- [12] L. Bass, P. Clements, and R. Kazman, *Software Architecture in Practice*, Second Edi. Addison Wesley, 2003.
- [13] F. Buschmann, R. Meunier, H. Rohnert, P. Sommerlad, and M. Stal, *Pattern-Oriented Software Architecture - Volume 1: A System of Patterns*. Wiley Publishing, 1996.
- [14] F. Bocchio, "Estudio Comparativo De Plataformas Cloud Computing Para Arquitecturas Soa," vol. 1, no. 5, p. 114, 2013.

- [15] S. Malik and D. H. Kim, "A comparison of RESTful vs. SOAP web services in actuator networks," *Int. Conf. Ubiquitous Futur. Networks, ICUFN*, pp. 753–755, 2017.
- [16] A. Arcuri, "RESTful API Automated Test Case Generation," *2017 IEEE Int. Conf. Softw. Qual. Reliab. Secur.*, pp. 9–20, 2017.
- [17] D. Booth *et al.*, "Web Services Architecture," *W3C Note*, vol. 22, no. February, pp. 1–98, 2004.
- [18] K. Gottschalk, S. Graham, H. Kreger, and J. Snell, "Introduction to Web services architecture," *IBM Syst. J.*, vol. 41, no. 2, pp. 170–177, 2002.
- [19] R. T. Fielding, "Architectural Styles and the Design of Network-based Software Architectures," *Building*, vol. 54, p. 162, 2000.
- [20] M. Zur Muehlen, J. Nickerson, and K. Swenson, "Developing web services choreography standards - The case of REST vs. SOAP," *Decis. Support Syst.*, vol. 40, no. 1 SPEC. ISS., pp. 9–29, 2005.
- [21] D. Fensel, F. M. Facca, E. Simperl, and I. Toma, *Semantic Web Services*, 1st ed. Springer Publishing Company, Incorporated, 2011.
- [22] C. Pautasso, O. Zimmermann, and F. Leymann, "Restful web services vs. 'big' web services: making the right architectural decision," *Proceeding 17th Int. Conf. World Wide Web*, pp. 805–814, 2008.
- [23] S. Patni, *Pro RESTful APIs*. 2017.
- [24] D. Flanagan, *JavaScript - The Definitive Guide*. 2011.
- [25] A. Q. Haviv, *MEAN Web Development*, vol. 1. 2014.
- [26] S. Holmes, *Getting MEAN with Mongo, Express, Angular, and Node*, 1st ed. Greenwich, CT, USA: Manning Publications Co., 2015.
- [27] K. Chodorow, *Mongo DB: The Definitive Guide*. 2013.
- [28] R. O. Obe and L. S. Hsu, *MongoDB in Action*. 2011.
- [29] E. Brown, *Web Development with Node and Express: Leveraging the JavaScript Stack*. O'Reilly Media, 2014.
- [30] E. Hahn, *Express in Action: Node Applications with Express and Its Companion Tools*, 1st ed. Greenwich, CT, USA: Manning Publications Co., 2015.
- [31] R. K. Soni, *Full Stack AngularJS for Java Developers*. 2017.
- [32] S. Davis, "Mastering MEAN: Introducing the MEAN stack," *IBM.com*, pp. 1–20, 2014.
- [33] M. Cantelon, M. Harter, T. J. Holowaychuk, and N. Rajlich, *Node.js in Action*, 1st ed. Greenwich, CT, USA: Manning Publications Co., 2013.
- [34] J. Bloch, *Effective Java (2Nd Edition) (The Java Series)*, 2nd ed. Upper Saddle River, NJ, USA: Prentice Hall PTR, 2008.
- [35] S. Daschner, *Architecting Modern Java EE Applications: Designing lightweight, business-oriented enterprise applications in the age of cloud, containers, and Java EE 8*. Packt Publishing, 2017.
- [36] F. Gutierrez, *Pro Spring Boot*. 2016.
- [37] J. Brisbin, O. Gierke, and G. Turnquist, "Spring Data REST - Reference Documentation," 2015. [Online]. Available: <https://docs.spring.io/spring-data/rest/docs/current/reference/html/>. [Accessed: 06-Apr-2018].
- [38] B. Muschko, *Gradle in Action*, 1st ed. Greenwich, CT, USA: Manning Publications Co., 2014.
- [39] S. Daschner, *Architecting Modern Java EE Applications: Designing lightweight, business-oriented enterprise applications in the age of cloud, containers, and Java EE 8*. Packt Publishing, 2017.
- [40] W3C, "HTML 5.2 Recommendation," 2017. [Online]. Available: <https://www.w3.org/TR/html52/>. [Accessed: 06-May-2018].
- [41] J. D. Gauchat, *El gran libro de HTML5. CSS3 y Javascript*, vol. 354, no. 9. 2012.
- [42] B. Noel and A. Chopra, "Introduction To Javaserer Pages," 2001.
- [43] "Siege Home," 2012. [Online]. Available: <https://www.joedog.org/siege-home/>. [Accessed: 14-May-2018].
- [44] "Apache HTTP server benchmarking tool." [Online]. Available: <https://httpd.apache.org/docs/2.4/programs/ab.html>. [Accessed: 14-May-2018].
- [45] "Apache JMeter - Apache JMeter™." [Online]. Available: <https://jmeter.apache.org/>. [Accessed: 13-May-2018].
- [46] "Prueba de rendimiento usando Jmeter." [Online]. Available: <https://www.guru99.com/jmeter-performance-testing.html>. [Accessed: 14-May-2018].
- [47] "Load Testing and Benchmarking With Siege." [Online]. Available: <https://www.serverwatch.com/tutorials/article.php/3936526/Load-Testing-and-Benchmarking-With-Siege.htm>. [Accessed: 14-May-2018].

Robot Asistente para Conducción Vehicular

Assistant robot for vehicles drivers

Aníbal Rubén Mantilla Guerra, Christian Alejandro Marín Castillo

Carrera de Ingeniería en Mecatrónica, Facultad de Ciencias de la Ingeniería e Industrias, Universidad UTE
Quito, Ecuador, ruben.mantilla@ute.edu.

Carrera de Ingeniería en Mecatrónica, Facultad de Ciencias de la Ingeniería e Industrias, Universidad UTE
Quito, Ecuador, chmarincastillo@hotmail.com

Resumen — Este documento presenta el proceso de desarrollo de un robot asistente personal para conductores de vehículos, considerando que una de las causas principales de los accidentes de tránsito es por distracción debida al uso de dispositivos tecnológicos como celulares y pantallas de video al momento de conducir. Con la determinación de las especificaciones técnicas para el diseño y construcción del equipo coadyuvante, fue posible usar de manera exitosa diversos métodos de programación, plataformas de código abierto, sistemas de control en lazo abierto, y herramientas de visión artificial OpenCV, en base a la norma ISO 25000. Se presentan los resultados de las pruebas con las cuales se validó el equipo.

Palabras Clave – Robot asistente, conducción vehicular, distracción, accidentes de tránsito

Abstract — This document presents the process of developing a personal assistant robot for vehicle drivers, considering that one of the main causes of traffic accidents is by cause of distraction due to the use of technological devices such as cell phones and video screens at the time of driving. With the determination of the technical specifications for the design and construction of the coadjutant equipment, it was possible to successfully use various programming methods, open source platforms, open loop control systems, and OpenCV artificial vision tools, based on the ISO 25000 standard. The results of the tests with which the equipment was validated are presented.

Keywords – Assistant robot, vehicular driving, distraction, traffic accidents

I. INTRODUCCIÓN

El Ecuador es un país en el cual, se ha incrementado de forma drástica tanto el comercio como el uso de automotores, en los últimos años, esto se ha visto reflejado en un mayor índice de siniestralidades como indican las cifras de la Agencia Nacional de Tránsito.

Para diciembre del 2017, se presenta un incremento tanto de las cifras de fallecidos, como

de heridos por accidentes de tránsito [1]; las causas de accidentes son múltiples, y son fundamentalmente las siguientes:

1. Conducir desatento a las condiciones de tránsito, por cualquiera de las siguientes acciones mientras se conduce: utilizar el celular o pantallas de video, ingerir comida, maquillaje o cualquier otro elemento distractor.
2. Conducir vehículo superando los límites máximos de velocidad.
3. No respetar las señales reglamentarias de tránsito (pare, ceda el paso, luz roja del semáforo, etc.).
4. No mantener la distancia prudencial con respecto al vehículo que le antecede.
5. Conducir bajo la influencia de alcohol, sustancias estupefacientes o psicotrópicas y/o medicamentos.

En la Fig. 1 se presentan las cifras de accidentes de tránsito en función de la causa para los años 2016 y 2017.

Principales causas de accidentes de tránsito



Figura. 1. Cifras totales de siniestralidad mostradas por su causa

Se podría disminuir el número de siniestros, si es que el conductor cuenta con un asistente vehicular para que mantenga su atención mientras conduce, haciendo que el conductor sea consciente del peligro que supone un lapsus de pocos segundos al distraerse del entorno por el cual transita

Tomando como referencia modelos de robots asistenciales como el “Robohon”, disponible en el mercado japonés, el cual integra funciones tanto de un robot humanoide de reducidas dimensiones, con las de un smartphone para integrarlos en un único dispositivo capaz de brindar ambas funcionalidades [2].

Otro ejemplo es el robot de la empresa Toyota denominado “Kirobo Mini” [3].

II. METODOLOGÍA

La metodología establecida para el desarrollo del equipo requiere que se determine primero los requerimientos y especificaciones:

- Requerimientos
 - Portable
 - Con Inteligencia Artificial
 - Monitoreo y toma de decisiones en tiempo real
 - Capacidad para generar diversos tipos de alerta
 - Peso no superior a 500 gramos
 - Autonomía mínima de 20 minutos
 - Fácil usabilidad
 - Fácil acoplamiento al vehículo
- Especificaciones técnicas generales
 - Tarjeta microprocesada a una velocidad mínima de 1.0 GHz
 - Visión Artificial
 - Control de servomotores
 - Pantalla LCD RGB 240 x 320
 - Batería capaz de suministrar 1,5 A de salida
 - Sistema de sonido
 - Dimensiones máximas 12 x 12 x 12 cm

Una vez establecidos los requerimientos y especificaciones técnicas del robot asistente personal, se presentan en esta sección los diferentes métodos utilizados para el desarrollo de cada subsistema.

A. MÉTODO DE VISIÓN ARTIFICIAL

Dado que el reconocimiento facial para un ser humano es una cualidad intrínseca, para una inteligencia artificial no lo es, ya que el software de implementación es un proceso muy complejo debido a los diferentes algoritmos y cálculos matemáticos que son requeridos para el sistema, es

por eso que se ha creado diferentes métodos de reconocimiento como detección de rostros, extracción de rasgos, identificación, incluso análisis de imágenes en 2 y 3 dimensiones debido al limitado rango de captura de imágenes que posee una cámara, resulta una ardua tarea para el programador.

Después de la segmentación se continua con la extracción de características de los objetos que van a permitir su clasificación. La selección de características es una etapa en la que se intenta identificar aquellas características que realmente aportan información cualitativa o relevante que son útiles para el proceso que se esté desarrollando. Las características de un objeto pueden ser internas o externas.

Las características externas permiten obtener información de la forma y el tamaño del objeto. Con el contorno se pueden hallar características morfológicas como: área proyectada, perímetro, longitud, volumen, diámetro, etc.

Los descriptores son un conjunto de números que se producen para describir una forma determinada. La forma puede no ser totalmente reconstruible a partir de los descriptores, pero estos deberían ser suficientes para que las diferentes formas sean discriminadas. Las regiones pueden describir propiedades basadas en límites de un objeto o pueden describir propiedades basadas en regiones [4]. Una imagen es una abstracción de la realidad en dos dimensiones, para que esta pueda interpretarse en un software se debe entender, como una serie de divisiones ínfimas representadas por píxeles, una vez entendido esto, se analiza dicho objeto capturado con los descriptores de región existentes como son los simples, topológicos, de textura y momentos geométricos que son los que se utilizan en el presente desarrollo del software.

Los descriptores geométricos que se han usado con más regularidad por varios años en reconocimiento de objetos y de visión por computadora han sido los momentos geométricos de Hu , los cuales son parte del procesamiento digital de imágenes, permiten el cálculo de la superficie de un objeto, incluso si ha tenido un cambio de tamaño o de giro con respecto a su posición inicial, a continuación, se explican sus funciones de cálculo. La descripción y los rasgos que describen un objeto puede explicarse mejor teniendo en cuenta que son derivados a partir del teorema de unicidad de Papoulis, que establece que: Si $f(x, y)$ es una función continua con valores diferentes de 0, solamente en una región finita del plano xy , un conjunto infinito de momentos de esa región m_{pq} , con $p, q > 0$ que están determinados de manera única por $f(x, y)$. Inversamente la secuencia de m_{pq} determina la función de $f(x, y)$ [5]. Además, si $f(x, y)$ es una función continua, positiva y acotada sobre una región finita del plano xy entonces el momento de orden $(p + q)$ de $f(x, y)$ se define como:

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy \quad (1)$$

La geometría de una región plana se basa en el tamaño, posición, orientación y forma. Todas estas medidas están relacionadas en los parámetros denominados momentos como se explicó a breves rasgos en el literal anterior.

Mediante los momentos invariantes de Hu , se consigue una descripción independiente a las posiciones, tamaños y ángulos del objetivo. Cuando se desea obtener un descriptor invariante respecto del tamaño y de la posición se puede utilizar momentos de Hu centrales normalizados de orden dos y tres [6]. Los momentos centrales de orden $(p+q)$ de una región en el caso continuo son invariantes a traslaciones y se define como:

$$u_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy \quad (2)$$

Donde:

$$\bar{x} = \frac{m_{10}}{m_{00}} \quad y \quad \bar{y} = \frac{m_{01}}{m_{00}} \quad (3)$$

De esta manera las invariantes de traslación se han transformado en momentos centrales, pero no se encuentran normalizados, es decir esta susceptible a cambios de escala. Para conseguir que sean invariantes a cambios de escala de debe dividir los momentos centrales por el tamaño del objeto de interés, es decir hacer una división simétrica del objeto focalizado. (Jain, 1989), como se muestra en la siguiente ecuación.

$$n_{pq} = \frac{u_{pq}}{v_{\infty}^{\gamma}} \quad (4)$$

Donde:

$$\gamma = \frac{p+q}{2} + 1 \quad (5)$$

Pero para el reconocimiento digital es preferible convertir las integrales en sumatorias para que exista un proceso continuo del proceso logarítmico tomando una cantidad para cada variable lanzada al infinito de esta manera.

$$u_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy \quad (6)$$

Así se puede explicar la funcionalidad de la ecuación utilizada en el desarrollo del software del reconocimiento facial y de ojos como se explica a continuación en el diseño del software.

La librería OpenCV utiliza los momentos de Hu para describir, caracterizar y cuantificar la forma de un objeto en una imagen, normalmente extrayendo la silueta o contorno del objeto, extrayéndolas en un vector de 7 elementos, que en este caso son los 7 momentos de Hu , siendo estos los que describen un patrón rostro o no rostro [8].

Para realizar los cálculos primero se debe realizar un cambio a escala de grises, es decir, se establecen los píxeles de la figura en fondo negro y el primer plano en blanco, para aplicar algún tipo de segmentación como se muestra en la Fig. 2:

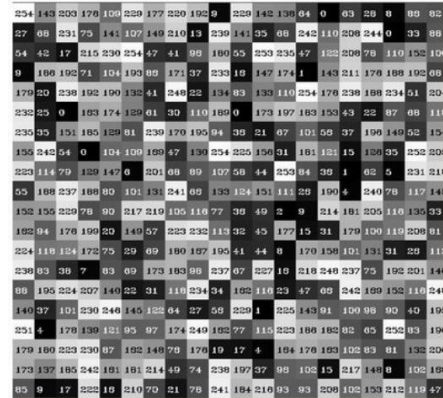


Figura. 2. Segmentación escala de grises

Una vez obtenido este procedimiento se puede mejorar la calidad de la imagen, la cual se conoce como ecualización. La ecualización del histograma, como se aprecia en la Fig. 3, trata de repartir los píxeles de forma que la luminancia de éstos se encuentre más distribuida. Consiguiendo así aumentar el contraste y distinguir mejor los distintos objetos que forman la imagen.



Figura. 3. Ecualización de escala de grises

OpenCV emplea un conjunto de bloques también conocidos como clasificadores, con el fin de reconocer distintas formas de objetos, éstas son conocidas como HaarLike Features; se las aprecia en la Fig. 4. Las características tipo Haar se definen sobre regiones rectangulares de una imagen en escala de grises. Una característica está formada por un número finito de rectángulos y su valor escalar consistirá en la suma de los píxeles de cada rectángulo suma aplicando un cierto factor de peso [8].

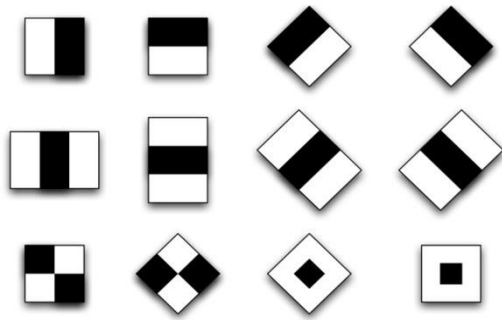


Figura 4. Haar-Like Features.

El valor de una característica se obtiene sumando todos los píxeles del rectángulo blanco y restándose todos los píxeles del rectángulo negro. Por ejemplo, la característica central de 3 rectángulos trataría de representar que en general la región de los ojos es más oscura que las regiones de alrededor.

Los descriptores que utiliza OpenCV, como quedó mencionado, son los de *Hu*, cada característica facial empleada (ojo izquierdo, ojo derecho) es descrita por $f(\theta_1, \theta_2, \theta_3, \theta_4)$, que en nuestro caso dentro de la programación se tiene una función: $for(x, y, w, h)$. De este modo se puede entender el algoritmo de detección usado en OpenCV, contenido dentro de los clasificadores Haar-Like Features para detectar el rostro y los ojos utilizando la función `cv2.CascadeClassifier` [8].

B. MÉTODO PARA EL DISEÑO DEL HARDWARE

Se hace uso de microcontroladores y plataformas de hardware considerando velocidad, memoria, facilidad de programación y uso.

B.1 TARJETA MICROCONTROLADORA ARDUINO

Esta se conecta a su vez con una plataforma central de procesamiento interactuando tanto a nivel de hardware como de software, para ejecutar el control de mecanismos que sirvan para crear interacción con el usuario.

B.2 TARJETA MICROPROCESADA RASPBERRY PI

Es un microcomputador de placa ensamblada de circuitos impresos de hardware libre el cual corre hasta una velocidad de 1,2 GHz, provisto de una memoria RAM, y una GPU, capaz de soportar todo un sistema operativo o actuar como un controlador

de software embebido, que se adapte a una función específica.

C. MÉTODO PARA EL DISEÑO DEL SOFTWARE

C.1 Software libre

Software libre es aquel con el que se puede hacer ciertas modificaciones o creación de cosas debido al consentimiento del autor, de aquí se puede destacar el movimiento Open Source (Open Source Initiative) que refleja más una política ligada a la ética y a la moral como libre distribución, códigos fuentes disponibles, licencias distribuidas como parches, sin discriminación de personas o grupos entre otros [9].

C.2 SO Raspbian

El sistema operativo instalado en la Raspberry pi está basado en Debian el cual es una distribución de Linux, optimizado para correr en el Hardware de la R-Pi, un sistema operativo es el conjunto de programas básicos y utilidades que hacen funcionar a la tarjeta electrónica y esta es capaz de ejecutar los diferentes programas que se utilizan en este proyecto como son Python, la librería Open CV y Arduino [10].

C.3 Python

Es un lenguaje de programación multiparadigma., administrado por Python Software Foundation. Posee una licencia de código abierto, denominada Python Software Foundation License [11].

C.4 Arduino IDE

El software de código abierto Arduino (IDE) hace que sea fácil escribir código y subirlo a la pizarra. Se ejecuta en Windows, Mac OS X y Linux. El entorno está escrito en Java y está basado en Processing y otro software de código abierto. Este software se puede usar con cualquier placa Arduino [12].

C.5 OpenCV

Viene de las siglas Open Source Computer Vision Library, es una librería abierta desarrollada por Intel en el año 1999, contiene alrededor de 500 funciones. Esta librería proporciona un alto nivel de funciones para el procesado de imágenes. Algunas de las características que permite OpenCV son operaciones básicas, procesamiento de imágenes, análisis estructural, análisis de movimiento, reconocimiento del modelo, reconstrucción 3D, calibración de cámara, etc.

Para el desarrollo del dispositivo tecnológico se considera como referente la norma ISO 25000 en base a los siguientes lineamientos: Compatibilidad, Usabilidad, Fiabilidad, Mantenibilidad y Portabilidad.

TABLA III. MÓDULO DE CÁMARA DE RASPERRY PI

Modelo	FOV160 ° 5 megapíxeles
Lente	¼ de pulgada
Sensor	OV5647
FOV	160 grados
Resolución	2592x1944
Dimensiones	22,5 x 24 x 9 mm

Los servomotores que mejor se adaptan al diseño del dispositivo, debido a su consumo energético y peso son idóneos para producir interacción con el usuario, sin intervenir en el tamaño del prototipo. Modelo Micro Servo 9g TowerPro; voltaje de operación de 3.0 hasta 7.2V; el torque máximo es de 1.8 Kg-cm; dimensiones 22 x 11.5 x 27 mm y el peso de 9 gramos.

Tomando en cuenta el consumo de la raspberry pi y de los diferentes elementos que la componen se toma por elegir un módulo propio de alimentación para la tarjeta procesadora la cual abastecerá de energía a todo el sistema y no sobrepasara las dimensiones que se requieren [16].

TABLA IV. CARACTERÍSTICAS DE LA BATERÍA

Modelo	RPI PowerPack V2.0
Capacidad de batería	3800 mAH
Corriente de descarga máxima	1.8 A
Voltaje de salida	5.1V ± 0.1V
Corriente / voltaje de carga estándar	1.0A / 5.0V

Para crear las alertas se utiliza un módulo amplificador de audio de 3 Watts. Modelo PAM8403; voltaje de operación de 2.5 – 5V; potencia de salida 2 x 3 W (5V/4 Ohm); dimensiones 20 x 20 x 3 mm y peso de 5 gramos.

En la Fig. 6, se muestra el sistema electrónico diseñado. La tarjeta principal de procesamiento es la encargada de llevar a cabo todos los procesos que se realizan, como la captura de imágenes, mediante una cámara propia para la Raspberry la cual integra un conector de tipo CSI (Camera Serial Interface), basado en un estándar definido por la MIPI (Mobile Industry Processor Interface), este tipo de conexión como se habló anteriormente permite una que se libere al CPU de esta tarea la cual queda centrada en la GPU (Graphics Processing Unit), que esta exclusivamente diseñada para el procesamiento y codificación de gráficos, dejando a la unidad central de procesamiento la toma de decisiones y diversas tareas que ejecuta el software. Al momento de comunicarse con la

unidad secundaria de control y procesamiento se realiza mediante USB por comunicación serial.

La unidad secundaria de control usa señales de PWM que son las encargadas de activar y mover los servos al momento de recibir la señales mediante comunicación serial que recibe de la unidad principal [17].

Al momento de cargar el sistema operativo y todos los programas que sobre este se ejecutan se utiliza una Micro SD de 16 GB clase 10, que es la recomendada por el fabricante; la cual tiene una tasa de transmisión de datos de hasta 10 MB/s, lo cual garantiza la correcta ejecución del software y del sistema operativo [18].

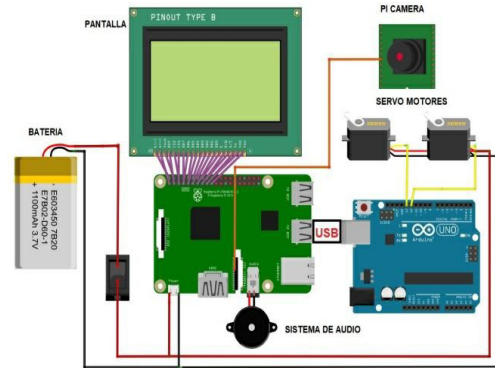


Figura 6. Sistema electrónico computacional diseñado

B. Diseño del software

Se puede indicar en un diagrama de bloques y un diagrama de flujo los diferentes procesos y funcionamiento del sistema, tomando en cuenta las diferentes fases del diseño y sus subprocesos [19].

- Fase 1.- se activa el sistema y se hace la captura de imágenes que ingresan por la cámara.
- Fase 2.- se activa el inicializador del trabajo que en este caso es la librería OpenCV, aquí con las instrucciones propias de la citada librería, se detecta y se marca la región de interés, la cual permite establecer lugares específicos de la imagen para codificar los datos de las características importantes de la zona de interés, el cual permite enfocar el rostro y los ojos para posteriormente realizar el conteo de la presencia o no de estos.
- Fase 3.- en este bloque se realizan las comparaciones de los aspectos físicos, cuando están los ojos abiertos o cerrados. Entonces se analiza al sujeto monitoreado para determinar los rasgos de cansancio o desconcentración.
- Fase 4.- al comparar los aspectos antes mencionados se ejecutan las alertas del asistente robótico; para emitir de forma pertinente dependiendo del grado de distracción; que va desde una señal de emergencia leve hasta la alerta de mayor severidad.

Todo sistema se debe especificar con un esquema para comprender sus procesos y los pasos a seguir del software. Es por ello, por lo que el diagrama de flujo es la base clave para el desarrollo del software del robot asistente, aquí se muestran todas las condiciones de cada variable, así como su funcionamiento para tomar las decisiones sobre las alertas emergentes, también los contadores para la emisión de las alertas y la detección de distracción del usuario.

La Fig. 7, muestra la arquitectura del software del sistema determinando el algoritmo de forma simple.

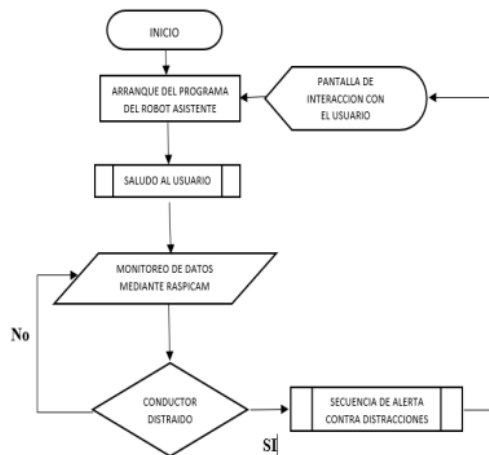


Figura 7. Arquitectura Básica del Software del sistema

IV. CONSTRUCCION

La Fig. 8, muestra todos los componentes electrónicos y de control con su respectivo nombre, los cuales van en el interior de la carcasa del robot.

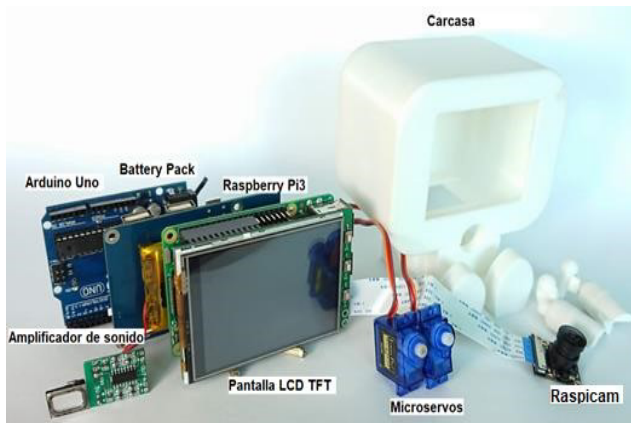


Figura 8. Vista en perspectiva de los componentes electrónicos del robot asistente

La Fig. 9, muestra los elementos que componen el robot asistente, para finalmente exhibir al mismo ensamblado, con todos los componentes de hardware y software. Uno de los requisitos planteados, fue el de darle una forma semejante o basada en la humana, para lo cual se diseñó una carcasa que posee características antropomórficas, esta medida se requiere para que sea más amigable al usuario.



Figura 9. Prototipo ensamblado con todas sus partes

V. ANALISIS DE RESULTADOS

Las pruebas ejecutadas se realizaron en un vehículo, bajo condiciones de máxima intensidad de luz solar durante el día, para que la cámara pueda captar de forma óptima el rostro del conductor y sus ojos. El robot asistencial es colocado frente al conductor con cualquier soporte para celulares, el cual se ubica sobre el parabrisas con una ventosa, lo que facilita su implementación en el vehículo como se muestra en la Fig. 10.



Figura 10. Prototipo colocado en el interior de la cabina para la realización del monitoreo, asistencia y emisión de las diferentes alertas.

A continuación, se presentan las pruebas que se realizaron para validar y verificar el correcto funcionamiento del prototipo, las pruebas son:

A. PRUEBA DE AUTONOMÍA ENERGÉTICA

Establecido que la batería debía tener una duración mínima de 20 minutos; se concluyó que esta suministra carga a todos los componentes en funcionamiento, por más de 21 minutos y 32 segundos, así el resultado fue totalmente favorable.

B. PRUEBA DE OPERACIÓN EN CONDICIONES DE ILUMINACIÓN CAMBIANTE

Se realiza una corta lista de control para verificar que el prototipo, funciona correctamente.

TABLA V. LISTA DE CONTROL DE RECONOCIMIENTO

Reconoce la morfología de rostro bajo las condiciones establecidas	x
Reconoce la morfología de ojos bajos las condiciones establecidas	x

C. PRUEBA DE NIVELES SONOROS

Cuando el vehículo está detenido con el motor apagado, el valor es de 39 dB, lo que indica que el ambiente es similar al que existe en una biblioteca, mientras que al momento de que el auto transita es de 69 dB, valor que se asemeja al que existe en una conversación.

D. PRUEBA DE FIABILIDAD EN RUTA

Tomando como criterios el número de aciertos y el número de alertas en diferentes tramos del recorrido, se realiza una estimación del porcentaje de aciertos usando una media aritmética entre el número de aciertos y el número de alertas, usando la siguiente ecuación:

$$\text{porcentaje}_{\text{aciertos}} = \frac{n_{\text{aciertos}}}{n_{\text{alertas}}} * 100\%$$

Al realizar el análisis de resultados se puede llegar a la deducción del funcionamiento que tiene el asistente personal, mostrados en la Fig. 11, se contabiliza un total de 18 alertas. Así mismo se muestran los límites de la intensidad de iluminación, por tanto los valores óptimos varían entre 400 luxes como mínimo y 1400 luxes como máximo; a pesar de estar en el rango de funcionamiento óptimo, aparecen dos alertas, las cuales pudieron presentarse por un exceso de vibraciones en el vehículo. El sistema opera mejor a condiciones bajas de iluminación que a valores elevados, esto ocurre porque se producen reflejos sobre el rostro, lo cual no permite detectar la morfología de la persona de manera correcta. Los puntos verdes son aciertos los rojos fallos.



Figura 11. Control de aciertos del robot asistencial relación a la iluminación de la cabina

Al momento de activarse las alertas se observa un incremento en 3 dB, con respecto al nivel sonoro que existía cuando se manejaba el vehículo, cuyo nivel sonoro fue de 69 dB, estos tres decibelios de sonido, descritos con anterioridad en las condiciones de prueba son suficientes para causar una respuesta positiva del conductor, éste ya es capaz de modificar su comportamiento y advertir sus distracciones. Según la escala decibélica este parámetro se encuentra en un nivel similar al que existe cuando hay tráfico en la ciudad.

Finalmente se puede constatar el funcionamiento del asistente robótico, al realizar una media aritmética utilizando la ecuación de porcentaje de aciertos y contabilizando el total de los mismos, esto es 13 aciertos sobre el total de 18 alarmas registradas. Se obtiene que la operabilidad total del robot asistente personal es de un 72.22%, es un valor aceptable si se toman en cuenta todas las condiciones de pruebas que se valoraron, las condiciones de la vía por la cual transita y el hecho de que es un primer prototipo operacional, el cual puede llegar a ser mejorado en varios aspectos electrónicos y de software.

VI. CONCLUSIÓN

(7)

Con la utilización efectiva de métodos y técnicas de ingeniería fue posible desarrollar un equipo tecnológico actual, que cumple satisfactoriamente la función de asistente en la conducción vehicular. Es compacto, ligero, de fácil uso, fiable, escalable, y con el potencial de poder ser utilizado en una gran cantidad de aplicaciones en las que se requiere visión artificial.

REFERENCIAS

- [1] "Agencia Nacional de Tránsito," 31 Diciembre 2017. [Online]. Available: <https://www.ant.gob.ec/>.
- [2] "Robohon," 31 Enero 2018. [Online]. Available: <https://robohon.com/>.
- [3] "Toyota Europa," 2018. [Online]. Available: <https://www.toyota-europe.com/world-of-toyota/articles-news-events/introducing-kirobomini>.
- [4] Ménard, Amy D. Contributors to Optimal Sexual Experiences. Diss. Université d'Ottawa/University of Ottawa, 2013.
- [5] Mahafza, Bassem R. Introduction to radar analysis. Chapman and Hall/CRC, 2017.

- [6] Narváez, Silva, Juan Esteban, and Ana María Zapata Herrera. "Reconstrucción de entornos 3d mediante un sistema de visión artificial estereoscópico, para aplicaciones en navegación robótica: etapa 1: implementación de un sistema de visión artificial estereoscópico para la estimación de las coordenadas 3-d de objetos presentes en un entorno bajo condiciones de luz controladas." (2015).
- [7] Lelescu, Dan, and Ankit K. Jain. "Feature based high resolution motion estimation from low resolution images captured using an array source." U.S. Patent No. 9,813,616. 7 Nov. 2017.
- [8] Open Source Computer Vision, "OpenCV," 10 07 2018. [Online]. Available: https://docs.opencv.org/3.4/d7/d8b/tutorial_py_face_detection.html. [9] Medrano, Carlos Molina. "El Software Libre y sus perspectivas hacia el futuro." (2014).
- [10] Raspberry foundation, "Raspberrypi.org," 23 07 2018. [Online]. Available: <https://www.raspberrypi.org/forums/viewtopic.php?t=192756>.
- [11] Open Source Computer Vision, "OpenCV," 10 07 2018. [Online]. Available: https://docs.opencv.org/3.4/d7/d8b/tutorial_py_face_detection.html.
- [12] Arduino, "Arduino," 31 01 2018. [Online]. Available: <https://www.arduino.cc/en/Main/Software>
- [13] López, Raquel Ramos. Identificación de la fuente en vídeos de dispositivos móviles. 2016.
- [14] Waveshare, "Waveshare," 25 05 2018. [Online]. Available: <https://www.waveshare.com/product/3.2inch-RPi-LCD-B.htm>
- [15] Villanueva, Moisés García, and Salvador Ramírez Zavala. "Integración de OpenCv en una Raspberry Pi: Sistema de Detección de Rostro vía WEB."
- [16] Sunfounder, "Sunfounder," 23 05 2018. [Online]. Available: <https://www.sunfounder.com/battery-power-pack.html> [17] S. Monk, Ejercicios prácticos con Raspberry Pi, Barcelona: Marcombo, 2017.
- [18] Laganier, Robert. OpenCV 3 Computer Vision Application Programming Cookbook. Packt Publishing Ltd, 2017.
- [19] Rondoy, Peña, and Elio Yoeni. "Implementación de un sistema informático de gestión de pagos en el instituto de educación superior tecnológico privado Santa Úrsula-Sullana; 2015." (2017).

Educación de Procesos de Negocio Basada en Generación Automática de Preguntas *Business Process Elicitation Based on Automatic Question Generation*

Gabriel Venegas, Carlos Cares

Departamento de Ingeniería Informática, Universidad Católica de Temuco, Temuco, Chile, gvenegas@inf.uct.cl .

Departamento Ciencias de la Computación e Informática, Universidad de la Frontera, Temuco, Chile,
carlos.cares@ceisufro.cl .

Resumen - Contexto: La educación de procesos de negocios es una actividad común tanto para la reingeniería de procesos de negocios como para la comprensión de dominios en la ingeniería de requerimientos. A pesar de esto, frecuentemente aparecen nuevas propuestas para realizar esta actividad, por lo que lo tratamos como un problema abierto. **Objetivo:** En este documento proponemos un método semi-automático para obtener procesos administrativos utilizando los principios de generación automática de preguntas. **Metodología:** Usamos un enfoque cualitativo multi-metodológico, primero realizamos un mapeo sistemático de literatura para mostrar los enfoques existentes, seguimos con un análisis conceptual para ilustrar sus fortalezas y debilidades, y finalmente usamos una prueba de concepto para demostrar la factibilidad de la solución propuesta. **Resultados:** Hemos obtenido una comparación de los enfoques existentes y una representación general del método que incluye una etapa específica de generación de preguntas. **Conclusión:** Se ha propuesto un método semi-automático para la obtención de procesos de negocios que muestra un método alternativo y prometedor a las clásicas entrevistas de educación de procesos.

*Palabras clave - Procesos de Negocio,
Generación Automática de Preguntas, Reingeniería
de Procesos, Ingeniería de Requerimientos.*

Abstract - Context: The business process elicitation is a common activity for both, business process re-engineering and domain understanding in requirements engineering. Despite this frequently appear novel proposals to develop them, thus it is an open problem. **Aim:** in this paper we propose a semi-automatic method for eliciting business process using automatic question generation principles. **Methodology:** We have used a qualitative multimethodological approach, first we use a systematic mapping literature review to show existing approaches, follow a conceptual analysis to illustrate their strengths

and weaknesses, and finally we use concept implementation to illustrate the feasibility of the proposed method. **Results:** We have obtained a comparison of existing approaches, and a general representation of the method including a specific stage of question generation. **Conclusion:** a semi-automatic method for business process elicitation has been proposed. It shows a promising alternative for classical interviews in this application domain.

Keywords - Business Process, Automatic Question Generation, Processes Re-engineering, Engineering Requirements.

I. INTRODUCCIÓN

Los procesos de negocio se definen como un conjunto de actividades coordinadas que apuntan a alcanzar un objetivo empresarial común [1]. Los modelos de procesos de negocio se consideran como una actividad importante que proporciona beneficios en las organizaciones. Estos procesos de negocio son documentos usados para: describir, mejorar y automatizar procesos existentes [2]. La necesidad de representar los procesos de negocio puede provenir de la necesidad de actualizar los procedimientos originalmente representados para una organización, si es el caso que esto ocurrió, o para entender y mejorar el quehacer de una organización cuyos procedimientos se definieron y han evolucionado a través del tiempo. Cuando el propósito de la representación de procesos es un cambio radical de los procedimientos, se habla de reingeniería de los procesos de negocio [3].

La educación del proceso de negocio significa lograr obtener una representación del proceso de negocio que se corresponda con la lógica organizacional en la cual ocurre. No es sólo un proceso de representar los dichos, sino muchas veces involucra consensuar creencias diferentes y formalizar prácticas sin previa

documentación en la organización. Esta obtención de procesos de negocio comienza con la comprensión de las actividades involucradas, en cuya representación se identifican entradas y salidas, flujos de actividades, puntos de decisión, actividades paralelas, entre otros elementos. Para efectuar esta educación existen básicamente enfoques tradicionales [4] y enfoques colaborativos [5],[6],[7]. Los enfoques tradicionales son las que aplican los analistas donde visitan a los interesados de todo tipo

(*stakeholders*) utilizando entrevistas para reunir información acerca de los procesos. Los enfoques colaborativos son los que se apoyan en herramientas de trabajo grupal que permite a los actores del proceso interactuar con otros para recolectar información de actividades realizadas. A pesar de la existencia de diversas propuestas, aún se logran apreciar desafíos en la obtención de procesos de negocio [5,11]. Algunos de esos desafíos son: comunicación inexacta, información imprecisa, lenta obtención del proceso de negocio, problema en la consistencia del modelo. Nosotros podemos agregar que, a pesar de la existencia de diferentes modelos de representación de negocios (UML, BPMN, entre otros), las entrevistas humano-humano sigue siendo la principal técnica de educación de flujos de trabajo administrativo.

Sin embargo, si consideramos que los procesos de negocio pueden ser descritos por metamodelos, es decir que es posible modelar en términos generales los procesos de negocio [8], entonces los tipos de preguntas y, por lo tanto sus respuestas, pueden establecerse como un conjunto de preguntas semánticamente acotadas, esto es, con tipos de respuestas pre-establecidas. En este ámbito, en el área de procesamiento de lenguaje natural existe toda una corriente de investigación que estudia la generación automática de preguntas (GAP) [9], y en particular existen propuestas para la generación de preguntas relativas a dominios de conocimiento específicos [10].

En este artículo hacemos una revisión de los métodos contemporáneos de obtención de procesos de negocio, presentamos nuestra propuesta de educación basada en generación automática de preguntas y su análisis y, finalmente, presentamos una comparación de los modelos existentes con el procedimiento propuesto.

En términos metodológicos hemos seguido un enfoque múltiple de metodologías cualitativas [11], primero aplicamos un mapeo

sistemático de literatura con el objetivo de encontrar las metodologías actuales, segundo realizamos un análisis conceptual para realizar la comparación de estas metodologías y finalmente utilizamos una prueba de concepto para mostrar que es posible generar un método basado en preguntas generadas automáticamente. En ingeniería de software enfoques como el análisis conceptual y prueba de concepto han sido habituales en la disciplina. Los resultados permiten no sólo tener un mapa resumido y comparable de las propuestas actuales de educación de procesos de negocio sino que, además, dan pie para profundizar en diferentes métodos para obtener procesos de negocio a partir de interacciones estructuradas y computarizadas. Este documento está organizado de la siguiente manera: la Sección 2 se discute sobre la obtención de procesos de negocio y las metodologías existentes; La Sección 3 presenta nuestra propuesta de educación de procesos en base a generación automática de preguntas. En la Sección 4 se presenta una comparación y, finalmente, en la Sección 5, resumimos nuestro trabajo presentado fortalezas, limitaciones y el trabajo futuro que tenemos planificado.

II. OBTENCIÓN DE PROCESOS DE NEGOCIO

Generalmente en la obtención de procesos de negocio se utiliza la técnica de la entrevista, esta se realiza la mayoría de las veces en el lugar de trabajo de los participantes del proceso [12][13][14]. Dependiendo del número de participantes los analistas realizan la obtención de la información en un lugar geográfico que se encuentra dentro de la misma empresa. Esta información el analista la registra en un papel, base de datos, audio o vídeo, lo que posteriormente requiere de una fase de análisis y representación.

Con el fin de lograr una representación real del proceso de negocio la realización de entrevistas se establece con diferentes participantes del proceso que, a su vez, podrían jugar diferentes roles en la organización. Sin embargo estos participantes pueden tener puntos de vistas limitados e incluso contradictorios del proceso en cuestión.

Otra complicación común que surge es que los participantes que se entrevistan no siempre cuentan con el tiempo necesario para participar. Esto provoca muchas veces que la entrevista sea delegada a personas que no tiene la experiencia necesaria sobre el proceso que se necesita obtener [15]. Estos participantes tienden a influenciar que el analista obtenga una interpretación incompleta y a veces errónea de los conceptos y formas de ejecución del proceso.

Bajo este escenario, los analistas se reúnen después de realizar las entrevistas, donde comparten la información obtenida. Dentro de esta reunión, se abordan dudas de la información y se construye el modelo del

proceso de negocio. La consolidación de la información sucede después de realizar las entrevistas, esto tiene como principal consecuencia que se identifiquen un alto número de inconsistencias fuera de la empresa objetivo, por lo que los analistas tendrán que planificar y ejecutar una nueva serie de entrevistas con el objetivo de resolver estas inconsistencias o incompletitudes.

Estos son, en general, los problemas que se abordan al intentar una educación de procesos de negocio y son los que la literatura técnica del área propone enfrentar de diferentes maneras. Existen por lo tanto variadas metodologías, las cuales podemos clasificar, en términos generales, entre tradicionales o clásicas y colaborativas.

En esta sección se mostrarán las principales metodologías presentadas en artículos recientes que permiten mejorar el proceso de la obtención de los procesos de negocio de una organización. Los artículos que presentan las metodologías descritas, fueron seleccionadas mediante el objetivo de identificar los estudios que se han realizado en los últimos 10 años, esto significó realizar un análisis sistemático de la literatura, en el cual los años seleccionados para la búsqueda fueron entre el 2008 hasta el 2018 (Agosto)..

La cadena de búsqueda utilizada como base para la obtención de los trabajos relevantes fue la siguiente:

Business and “Process” and Elicitation. Esta búsqueda se realizó en las siguientes fuentes de datos: IEEE, Scopus, Springer y Google scholar.

Para filtrar estudios se utilizaron los siguientes criterios: (i) *inclusión*: artículos, publicaciones en revistas, congresos y workshops que describen metodologías para la educación de flujos de trabajo. Incluso aunque solo tenga acceso al resumen del mismo, y (ii) *exclusión*: *descartando* artículos que no reportan metodologías para la educación de flujos de trabajo y artículos que se centran en el diseño de un flujo de trabajo, pero que no expresa una metodología para ello. De este modo obtuvimos 5 artículos (2009-2015) dedicados a explicar una metodología para la obtención de procesos de negocio, los años investigados indican que se trata de un problema que, pese a lo clásico, no deja de recibir atención en cuanto a oportunidades de mejora.

A. Metodología PASW

PASW [16] En esta propuesta la idea principal es involucrar a los empleados de una organización que se someten a una reingeniería de procesos de negocio en el área. Esta participación debe realizarse en seis fases consecutivas: aprendizaje, proceso de obtención, alternativas y soluciones, evaluación de opciones, implementación del flujo de trabajo y mantenimiento del proceso (ver Figura 1).

En la fase de aprendizaje: se considera que deben definirse los objetivos y entrenar a los participantes al proceso que serán sometidos. En la fase proceso de obtención se obtiene, documenta y detecta problemas del proceso. En la fase alternativas y soluciones, se generan y proponen soluciones específicas que estén vinculadas con los niveles estratégicos, tácticos y operacionales de la empresa. En la fase evaluación de opciones, se generan los criterios de designación que permitirán evaluar alternativas para posteriormente seleccionar la mejor alternativa de solución al proceso. En la fase implementación del flujo de trabajo, se realiza la implementación de las soluciones generando nuevos flujos, y finalmente la fase de mantenimiento del proceso, se encarga de realizar mantenencias a los flujos por la vía de iterar en conversaciones para mejorar los flujos de trabajo.

B. Metodología ISEA

ISEA se basa en los objetivos tradicionales de BPM, es decir de representar los aspectos relevantes para el negocio, dejando fuera los elementos complejos de la realidad. Para obtener los procedimientos utiliza un enfoque participativo, basado en una herramienta que da soporte al propio enfoque ISEA generando como resultado un proceso básico similar al BPMN acompañado del juego de rol (ver Figura 2), durante el juego, las actividades y enlaces pueden ser eliminado o modificado solo por el participante que los crea. Además, ISEA suministra usuarios con un lenguaje y actividades participativas que permiten obtener y mejorar las perspectivas complementarias de los procesos de negocio [17].

Este enfoque participativo de modelado del usuario final, permite a las partes interesadas en un proceso de negocio a colaborar juntas, de manera sencilla para comunicar y mejorar la obtención del proceso. Sus autores destacan que el procedimiento sugerido es preciso y estable. Este enfoque abarca la perspectiva organizacional de los procesos de negocio y explota la información compilada durante la obtención de la perspectiva organizacional. Y, adicionalmente, hace mención a una perspectiva interactiva que permite a los usuarios revisar la coherencia dentro del proceso.

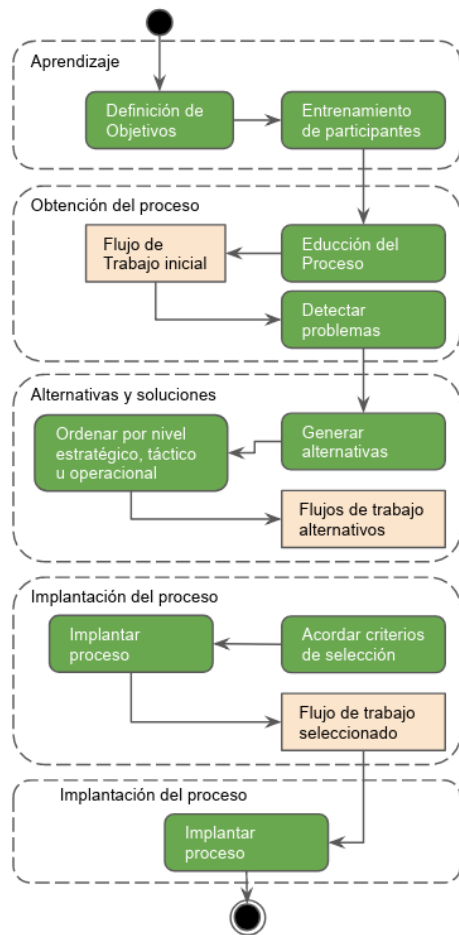


Figure 1. Fases y Tareas de PASW [16].

C. Metodología colaborativa/contextual

Los autores [14] proponen un modelo de contexto que se utiliza como base para la gestión de herramienta de contexto para apoyar la obtención de proceso de negocios. la idea es que la información obtenida sea comparada con información planteada anteriormente en otras entrevistas.

El enfoque se basa en diagramas en notación BPMN que el analista realiza durante la entrevista a lo stakeholders, siendo los mismo diagramas agregados con información contextual y compartidos entre los analistas a cargo del proyecto de modelado de procesos de negocio (ver figura 4). Una característica de este enfoque es que la información recopilada sobre el proceso debe estar asociada con su contexto dentro de la entrevista, de modo que los analistas puedan tener acceso únicamente a la información que sea relevante para la obtención que se está realizando [14].

El acceso a esa información permite que el analista tenga una comprensión más amplia del proceso que se está esbozando, lo que permite la facilidad de la identificación y resolución de conflictos. Además de contar con acceso a información relevante durante la entrevista tiene un impacto directo a la integridad y consistencia del modelo obtenido.

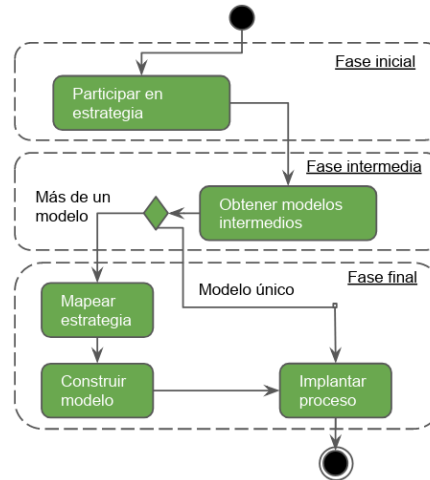


Figure 2. Fases y Tareas de ISEA [17].

D. Metodología colaborativa móvil

Este enfoque (ver figura 4) se basa en la utilización de la colaboración mediante una aplicación móvil, que permite ir generando e bosquejo del proceso de negocio, mientras se realiza la entrevista, esta herramienta móvil permite la colaboración de distintos analistas que se encuentren realizando entrevistas en la organización, esta colaboración es con ayuda de la aplicación móvil que permite crear una conexión ad-hoc, donde todos tienen acceso al bosquejo que se va realizando, esto permite que exista retroalimentación en tiempo real de las actividades, roles, nodos de decisiones en el bosquejo del proceso. Si bien pareciera que contar con una aplicación móvil que permita el diagramado del proceso fuese una ventaja, se puede convertir en una debilidad, ya que el ritmo de la entrevista es mucho más rápido, lo que provoca que la aplicación visual de dibujo sea lenta [18],[19].

E. Metodología Juego de roles virtuales

Los autores [20], proponen un método alternativo para obtener información en el proceso de negocio (ver figura 5), sin la necesidad de que los participantes tengan que explicar verbalmente todos los pasos involucrados en

una tarea grande. Esto se realiza utilizando el mundo virtual 3D que refleja el entorno de ejecución del proceso interesado. Aquí el usuario puede recorrer el mundo diseñado y representar roles en cada paso que normalmente realizaría en el proceso [20].

Esta herramienta está construida con el enfoque de juego de roles que utiliza la gramática de modelado de

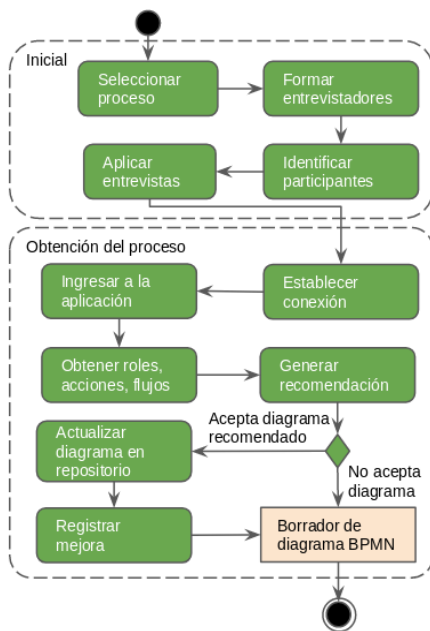


Figure 3. Fases y Tareas de enfoque colaborativo contextual [14]

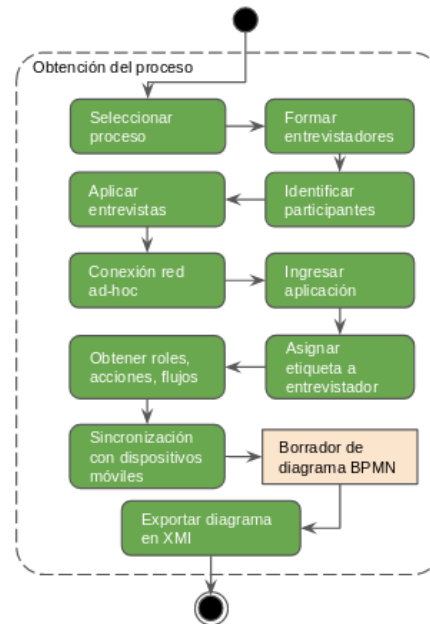


Figure 4. Fases y Tareas de enfoque colaborativo móvil [18],[19]

Gestión de procesos de negocios orientada a temas (S-BPM).

III. UNA PROPUESTA DE EDUCACIÓN BASADA EN GENERACIÓN AUTOMÁTICA DE PREGUNTAS

Si bien, dentro de la comunidad científica existen métodos y herramientas que permiten mejorar la obtención de los procesos de negocio, este proceso sigue siendo tedioso, ya que las propuestas describen una metodología enfocada en la colaboración, esto implica que las herramientas implementadas necesitan que en la obtención de procesos de negocio varios analistas están en la organización en la que se necesita obtener los procesos a capturar, volviéndolo un proceso lento y que impacta en las horas hombre de las personas involucradas en el proceso [18], [21]. Por este motivo, se propone un nuevo enfoque que consiste en desarrollar e implementar un método sencillo de obtención de procesos de negocio mediante un cuestionario online y fácil de usar, que consta de preguntas cerradas que permitirá obtener y analizar los procesos de negocio. También consta con un segundo cuestionario que permitirá resolver los conflictos encontrados al aplicar el primer cuestionario, esto lo realiza de forma automática/semi-automática.

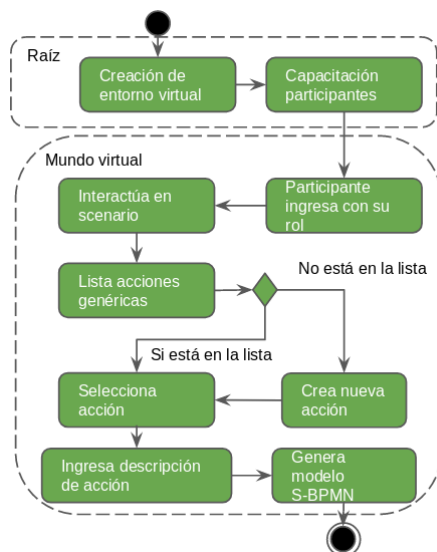


Figure 5. Fases y Tareas de enfoque de juego de rol virtual [20].

Como se muestra en la Figura 6, el enfoque consta de 3 participantes, los que se describen a continuación:

1) *Analista*: Es la persona que está a cargo de realizar la obtención del proceso de negocio

2) *Participante*: Son las personas identificadas que participan en el proceso que se pretende obtener.

3) *SEW*: Es el sistema automático que se encargará de obtener el proceso de negocio utilizando las respuestas del cuestionario generado.

A continuación se describen las actividades que presenta el enfoque propuesto en la figura 6:

A. *SETUP*:

En esta actividad, se incorporan todas las tareas relacionadas al inicio del proceso, es decir reuniones con los stakeholders, identificación del proceso a analizar.

B. *Genera cuestionario online*:

Esta actividad es automática, ya que el propio sistema SEW es el que se encarga de generar el enlace con las

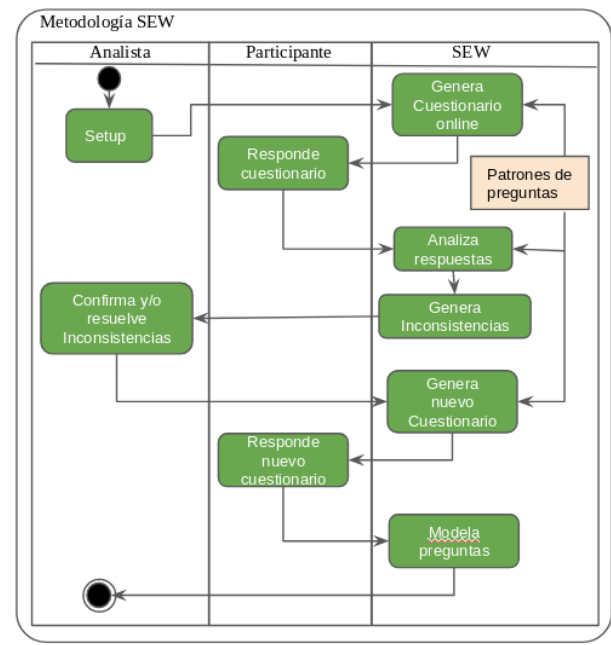


Figure 6. Enfoque SEW propuesto

Preguntas del primer cuestionario online que debe responder los participantes identificados en el SETUP.

C. *Responder cuestionario Online*:

Esta actividad la realiza los participantes identificados del proceso. Aquí ellos responden las preguntas cerradas con las que cuenta el formulario. Estas preguntas están debidamente identificadas para obtener un proceso de negocio.

D. *Analiza las respuestas*:

Una vez que todos los participantes responden el cuestionario, el sistema procederá a analizar cada una de las respuestas generadas por los participantes.

E. *Genera inconsistencia*:

Esta actividad la realiza el sistema SEW, aquí el sistema automáticamente se encargará de encontrar problemas del proceso, falta de decisiones, roles duplicados, tareas que realiza personas en paralelo.

F. *Confirma/resuelve inconsistencias*:

Esta actividad la realiza el analista, una vez que el sistema genera las inconsistencias encontradas, el analista procede a revisar para poder descartar ciertas anomalías encontradas, Un caso de ejemplo puede ser que dos tareas sean la misma pero los participantes la identificaron con distintos nombres.

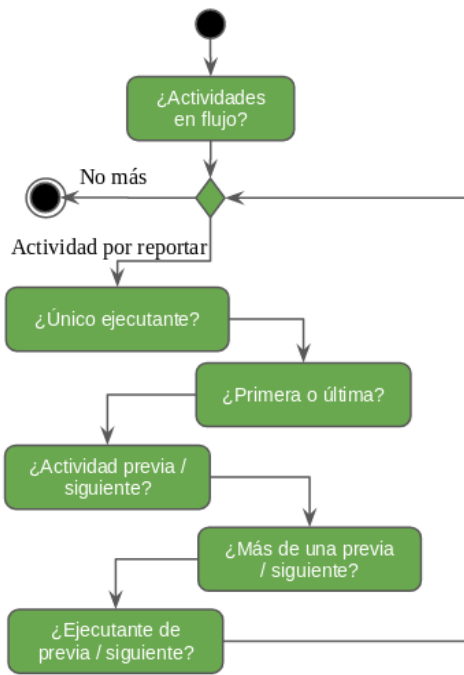


Figure 7. Algoritmo de generación automática de preguntas (GAP) para la educación de procesos de negocio.

G. Genera nuevas preguntas:

Esta actividad la realiza el sistema SEW, una vez que se hayan confirmado las inconsistencias encontradas. Para ello se genera un nuevo cuestionario, que permite resolver las inconsistencias encontradas en la actividad anterior.

H. Responde preguntas:

Aquí nuevamente el participante debe responder las preguntas que permitan resolver las inconsistencias. La forma en que son generadas las preguntas aparece en la Figura 7.

I. Modelo preguntas:

Aquí se pretende que el sistema pueda generar un mecanismo en el que se pueda entregar un modelo visual del proceso obtenido.

IV. COMPARACIÓN

En esta sección se muestra la comparación de los enfoques presentados en la sección 2. Se trata de una comparación de atributos de calidad, que no necesariamente representan una bondad *a priori*, sino que se valora con un objetivo específico, en

base al consumo de recursos en la obtención del proceso de cada enfoque analizado. Hemos escogido un enfoque multimetodológico, primero usamos un mapeo sistemático de literatura para mostrar los enfoques existentes, seguimos con un análisis conceptual para ilustrar sus fortalezas y debilidades (Ver tabla I).

TABLA I:
Valoración de los criterios de comparación para orientación a optimizar recursos

	Criterios
Ventajas	Generación automática de preguntas
	Soporte de aplicación visual
	Soporte automático de conflictos (Inconsistencias)
	Recomendaciones automáticas
Desventajas	Entrevista presencial
	Soporte de conversación y coordinación
	Soporta varios analistas
	Capacitación previa a los stakeholder (Participantes)
	Soporte en tiempo real de modelos de proceso

Nota: Tabla que agrupa los criterios de características de la comparación de los enfoques en ventajas y desventajas en base al consumo de recursos. Fuente: Elaboración propia.

A continuación se describen cada una de las características utilizadas en la tabla de comparación:

1) Entrevista presencial: Esto indica que la técnica utilizada para obtener el proceso se aplica de forma presencial entre el analista y stakeholder.

2) Automático en preguntas acotadas: Indica si el enfoque propuesto realiza la obtención del proceso utilizando una generación automática de preguntas.

3) Soporta varios analistas: Indica que el enfoque permite que una cantidad de N+1 analista participe en el proceso de obtención.

4) Soporte de aplicación visual en entrevista: Indica si existe una herramienta que apoye la obtención del proceso de negocio de forma visual.

5) Resuelve inconsistencias: Indica si el enfoque tiene una forma de encontrar y resolver conflictos que se encuentren en el proceso.

6) Capacitación previa: Indica que antes de realizar la obtención del proceso se necesita de una inducción a los stakeholders para participar de la obtención.

7) Soporte de conversación y coordinación: Indica si el enfoque permite que los analistas y stakeholders que participan de forma colaborativa dentro del proceso.

8) Recomendaciones automáticas: Indica si existen sugerencias que aparecen de forma automática al momento de obtener el proceso de negocio.

9) Soporte automático de conflictos: Indica si dentro del enfoque hay buscador de inconsistencias que muestra los conflictos encontrados en el proceso de forma automática.

La Tabla II muestra una síntesis de características de cada una de las metodologías revisadas en la sección anterior: PASW, ISEA, colaborativa móvil (CM), colaborativa contextual (CC) y juego de roles virtuales (JR), comparada con la propuesta de este trabajo (SEW).

Criterios para completar el cuadro de comparación:

- 1) S: Cumple la característica
- 2) N: No cumple la característica
- 3) I: No aplica

Dentro de la comparación realizada resaltan los enfoques colaborativos como los más utilizados con ayuda de herramientas móviles [18],[19] y medios contextuales [14], que mejoran la movilidad, visualización del modelo en tiempo real, consolidan la información obtenida mediante un repositorio global y realizan el intercambio de información en tiempo real, algo que destaca del enfoque colaborativo contextual es que genera recomendaciones en tiempo real de procesos similares que se encuentren dentro de su base de datos. A su vez las debilidades más notorias son la limitación de conexión para acceder al repositorio debido que pueden existir zonas geográficas o de infraestructura que no posean o permiten conexión a internet. Otro punto débil es la aplicación visual, ya que es muy lenta para seguir el ritmo de la entrevista.

Otro enfoque interesante es el propuesto por [20] juego de roles virtuales, dentro de las ventajas del enfoque tenemos: una mejor categorización de proceso, se recuerda con mayor facilidad la estructura del proceso, requiere poca asistencia externa, pero si necesita una capacitación previa. Si bien las tecnologías 3D han avanzado, no se garantiza que sea fácil desarrollar en un tiempo determinado el escenario del juego virtual, ya que cada proceso tiene distintos niveles de detalle, otro problema es el tiempo que se les puede permitir aún participante se pueda mover por el mundo virtual, por lo que

nos encontramos con el problema del tiempo de obtención del proceso.

El cuadro comparativo refleja un mayor uso de la entrevista como técnica en obtención del proceso, además indica que 4 de 5 enfoques revisados interactúan en la entrevista de forma humano a humano, esto demuestra que nuestro método SEW es prometedor, ya que proponemos capturar el proceso utilizando un cuestionario online automático con preguntas acotadas. Si bien, lo anterior es importante, debemos mencionar que nuestra propuesta tiene una debilidad en la coordinación entre los analistas, que es fundamental para obtener un proceso de negocio más robusto.

El análisis demuestra que nuestro enfoque SEW propone considerar un soporte de inconsistencia de forma automática, es decir mostrar conflictos que encuentre al momento de obtener el proceso. Esto promete resolver los conflictos de forma más eficiente.

TABLA II:
Cuadro comparativo de los enfoques

Características/Enfoques	PASW	ISEA	CM	CC	JR	SEW
Entrevista presencial	S	S	S	S	S	N
Automático en preguntas acotadas	N	N	N	N	N	S
Soporta varios analistas	S	S		S	N	I
Soporte de aplicación visual en entrevista	N	S	S	S	S	N
Capacitación previa al stakeholders al someterse al proceso de obtención	N	N	N	N	S	N
Soporte de conversaciones y coordinación	S	S	S	S	S	N
Recomendaciones Automáticas en el proceso de obtención.	N	N	S	N	N	N
Soporte automático de conflictos (Inconsistencias)	I	N	N	N	N	S

Nota: Tabla comparativa de las características encontradas en los enfoques revisados. Fuente: Elaboración propia para mostrar las características de los enfoques.

Otro punto importante que demuestra el cuadro es que existe una nicho de investigación que no se ha trabajado lo suficiente ya que solo 1 de los 5 enfoques revisado genera recomendaciones automáticas en la obtención de procesos de negocio.

V. CONCLUSIONES Y TRABAJOS FUTUROS

En este trabajo hemos mostrado que la educación de procesos de negocio sigue siendo un área de preocupación tanto desde la perspectiva de reingeniería de procesos de negocio como de ingeniería de requerimientos. Hemos realizado un mapeo sistemático de literatura para obtener procesos de negocio estos últimos 10 años y hemos obtenido 5 metodologías con este propósito. Todas ellas tienen como técnica relevante y necesaria, las entrevistas humano-humano entre analistas y stakeholders. Nosotros hemos propuesto una metodología que apoya el proceso a partir de un algoritmo automático de generación de preguntas y un algoritmo semiautomático para identificar y resolver inconsistencias de la información recopilada automáticamente. Adicionalmente, estos procesos los hemos representado en un marco de actividades general con el fin de visualizar todas las fases de la metodología propuesta.

Adicionalmente, hemos realizado una comparación de la metodología propuesta con las otras metodologías contemporáneas apareciendo en nuestro cuadro de evaluación áreas de claras ventajas de nuestra propuesta. Así por ejemplo queda en evidencia que ninguna de las otras propuesta disminuye una de las variables más críticas de este proceso, cual es la interacción humano-humano.

Dado lo anterior el aporte real de nuestra propuesta es un enfoque que permite disminuir la interacción entre el analista y los stakeholder obteniendo el proceso mediante preguntas acotadas y de forma automática, esto es importante ya que no solo ahorrará tiempo a los analistas y stakeholders, si no también permitirá disminuir el tiempo de obtención del proceso de negocio haciéndolo un proceso más eficiente.

Dentro de las limitaciones, y es donde enfocamos nuestro trabajo futuro, es validar casos reales y recoger información no sólo de la factibilidad técnica sino de la eficiencia, habilitando la posibilidad de realizar simulaciones y, obtener así, una comparación

cuantitativa que nos permita aproximar una magnitud de la mejora en la educación de procesos de negocio llevados a cabo con nuestra metodología.

REFERENCIAS

- [1] M. CHINOSI AND A. TROMBETTA, "BPMN: AN INTRODUCTION TO THE STANDARD," *COMPUT. STAND. INTERFACES*, VOL. 34, NO. 1, PP. 124–134, 2012.
- [2] P. ANTUNES, V. HERSKOVIC, S. F. OCHOA, AND J. A. PINO, "MODELING HIGHLY COLLABORATIVE PROCESSES," IN *PROCEEDINGS OF THE 2013 IEEE 17TH INTERNATIONAL CONFERENCE ON COMPUTER SUPPORTED COOPERATIVE WORK IN DESIGN (CSCWD)*, 2013.
- [3] R. ARCHER AND P. BOWKER, "BPR CONSULTING: AN EVALUATION OF THE METHODS EMPLOYED," *BUSINESS PROCESS MANAGEMENT JOURNAL*, VOL. 1, NO. 2, PP. 28–46, 1995.
- [4] A. SHARP AND P. MCDERMOTT, *WORKFLOW MODELING: TOOLS FOR PROCESS IMPROVEMENT AND APPLICATIONS DEVELOPMENT*. ARTECH HOUSE, 2009.
- [5] R. M. DE FREITAS, M. R. S. BORGES, F. M. SANTORO, AND J. A. PINO, "GROUPWARE SUPPORT FOR COOPERATIVE PROCESS ELICITATION," IN *LECTURE NOTES IN COMPUTER SCIENCE*, 2003, PP. 232–246.
- [6] J. C. DE A. R. GONCALVES, F. M. SANTORO, AND F. A. BAIAO, "A CASE STUDY ON DESIGNING BUSINESS PROCESSES BASED ON COLLABORATIVE AND MINING APPROACHES," IN *THE 2010 14TH INTERNATIONAL CONFERENCE ON COMPUTER SUPPORTED COOPERATIVE WORK IN DESIGN*, 2010.
- [7] F. M. SANTORO, M. R. S. BORGES, AND J. A. PINO, "CEPE: COOPERATIVE EDITOR FOR PROCESSES ELICITATION," IN *PROCEEDINGS OF THE 33RD ANNUAL HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES*.
- [8] S. SMIRNOV, *BUSINESS PROCESS MODEL ABSTRACTION: THEORY AND PRACTICE*. UNIVERSITÄTSVERLAG POTSDAM, 2010.
- [9] M. DIVATE AND A. SALGAONKAR, "AUTOMATIC QUESTION GENERATION APPROACHES AND EVALUATION TECHNIQUES," *CURR. SCI.*, VOL. 113, NO. 09, P. 1683, 2017.
- [10] G. KURDI, B. PARSIA, AND U. SATTLER, "AN EXPERIMENTAL EVALUATION OF AUTOMATICALLY GENERATED MULTIPLE CHOICE QUESTIONS FROM ONTOLOGIES," IN *LECTURE NOTES IN COMPUTER SCIENCE*, 2017, PP. 24–39.
- [11] R. L. GLASS, I. VESSEY, AND V. RAMESH, "RESEARCH IN SOFTWARE ENGINEERING: AN ANALYSIS OF THE LITERATURE," *INFORMATION AND SOFTWARE TECHNOLOGY*, VOL. 44, NO. 8, PP. 491–506, 2002.
- [12] A. GROSSKOPF, J. EDELMAN, AND M. WESKE, "TANGIBLE BUSINESS PROCESS MODELING – METHODOLOGY AND EXPERIMENT DESIGN," IN *LECTURE NOTES IN BUSINESS INFORMATION PROCESSING*, 2010, PP. 489–500.
- [13] F. M. SANTORO, M. R. S. BORGES, AND J. A. PINO, "ACQUIRING KNOWLEDGE ON BUSINESS PROCESSES FROM STAKEHOLDERS' STORIES," *ADVANCED ENGINEERING INFORMATICS*, VOL. 24, NO. 2, PP. 138–148, 2010.
- [14] D. M. SILVA, R. M. DE ARAUJO, F. M. SANTORO, AND G. A. P. PASCUAL, "DEFINING CONTEXT IN A BUSINESS PROCESS COLLABORATIVE ELICITATION APPROACH," IN *PROCEEDINGS OF THE 2012 IEEE 16TH INTERNATIONAL CONFERENCE ON COMPUTER SUPPORTED COOPERATIVE WORK IN DESIGN (CSCWD)*, 2012.
- [15] J. COUGHLAN, M. LYCETT, AND R. D. MACREDIE, "COMMUNICATION ISSUES IN REQUIREMENTS ELICITATION: A CONTENT ANALYSIS OF STAKEHOLDER EXPERIENCES," *INFORMATION AND SOFTWARE TECHNOLOGY*, VOL. 45, NO. 8, PP. 525–537, 2003.
- [16] "IMPROVING THE EFFECTIVENESS OF BUSINESS PROCESS DEVELOPMENT THROUGH COLLABORATION ENGINEERING: A METHOD FOR PROCESS ELICITATION," IN *2009 42ND HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES*, 2009.
- [17] A. FRONT, D. RIEU, M. SANTORUM, AND F. MOVAHEDIAN, "A PARTICIPATIVE END-USER METHOD FOR MULTI-PERSPECTIVE BUSINESS PROCESS ELICITATION AND IMPROVEMENT," *SOFTWARE & SYSTEMS MODELING*, VOL. 16, NO. 3, PP. 691–714, 2015.

- [18] N. BALOIAN *ET AL.*, “A COLLABORATIVE MOBILE APPROACH FOR BUSINESS PROCESS ELICITATION,” IN *PROCEEDINGS OF THE 2011 15TH INTERNATIONAL CONFERENCE ON COMPUTER SUPPORTED COOPERATIVE WORK IN DESIGN (CSCWD)*, 2011.
- [19] N. BALOIAN, J. A. PINO, C. REVECO, AND G. ZURITA, “MOBILE COLLABORATION FOR BUSINESS PROCESS ELICITATION FROM AN AGILE DEVELOPMENT METHODOLOGY VIEWPOINT,” IN *2013 IEEE 10TH INTERNATIONAL CONFERENCE ON E-BUSINESS ENGINEERING*, 2013.
- [20] J. HARMAN, R. BROWN, D. JOHNSON, S. RINDERLE-MA, AND U. KANNENGIESSER, “VIRTUAL BUSINESS ROLE-PLAY: LEVERAGING FAMILIAR ENVIRONMENTS TO PRIME STAKEHOLDER MEMORY DURING PROCESS ELICITATION,” IN *LECTURE NOTES IN COMPUTER SCIENCE*, 2015, pp. 166–180.
- [21] M. A. CHIARELLO, MARIA CLÁUDIA F, AND ADOLFO GUSTAVO S, “AN APPROACH OF SOFTWARE REQUIREMENTS ELICITATION BASED ON THE MODEL AND NOTATION BUSINESS PROCESS (BPMN),” *LECTURE NOTES ON SOFTWARE ENGINEERING*, pp. 65–70, 2014.

IMPACTO EN LA AGILIDAD DE LA ANALÍTICA DE DATOS USANDO TECNOLOGÍA IN-MEMORY. IMPACT ON THE AGILITY OF DATA ANALYTICS USING IN- MEMORY TECHNOLOGY.

Mario Raúl Morales Morales, Santiago Morales Cardoso, Milton Giovanni Moncayo Unda, Bryan Agustin Saltos Saá, Mónica Marianela Alvear Merizalde

Carrera de Ingeniería Informática, Facultad de Inge Cincias Físicas y Matemática, Universidad Central del Ecuador, mmoralesm@uce.edu.ec

Carrera de Ingeniería Informática, Facultad de Inge Cincias Físicas y Matemática, Universidad Central del Ecuador, smorales@uce.edu.ec

Carrera de Ingeniería Informática, Facultad de Inge Cincias Físicas y Matemática, Universidad Central del Ecuador, mmoncayo@uce.edu.ec

Carrera de Ingeniería Informática, Facultad de Inge Cincias Físicas y Matemática, Universidad Central del Ecuador, basaltos@uce.edu.ec

Carrera de Ingeniería Informática, Facultad de Inge Cincias Físicas y Matemática, Universidad Central del Ecuador, mmalvear@uce.edu.ec

Resumen. - A través de una metodología exploratoria este trabajo ha analizado los factores de eficiencia que han propiciado una mayor agilidad en los sistemas de Inteligencia de Negocios basados en uso de tecnología In-Memory (IM). Para ello se estudiaron las características de bases de datos NoSQL y particularmente las propiedades del almacenamiento columnar; algunas consideraciones de eficiencia tales como el no acceder a disco, la gestión de rendimiento de datos en memoria, las técnicas de compresión, las operaciones de barrido y el aprovechamiento del multiprocesamiento y paralelismo. Posteriormente se revisaron casos de estudio reales donde estas tecnologías han sido aplicadas y cuáles han sido los resultados conseguidos. Todas estas características favorecen la agilidad en los proyectos de Inteligencia de Negocios, logrando mejores rendimientos incluso sobre grandes volúmenes de datos a un menor costo total de propiedad.

Palabras clave: *In-Memory, Base de datos columnar, Inteligencia de negocios, Tempo real, Rendimiento.*

Abstract. - Through an exploratory methodology, this work has analyzed the efficiency factors that have led to greater agility in Business Intelligence systems based on the use of In-Memory (IM) technology. To this end, the characteristics of NoSQL databases and particularly the properties of columnar storage were studied; Some efficiency considerations such as not

accessing disk, memory data performance management, compression techniques, scanning operations and the use of multiprocessing and parallelism. Subsequently, real case studies were reviewed where these technologies have been applied and what the results have been achieved. All these characteristics favor agility in Business Intelligence projects, achieving better returns even on large volumes of data at a lower total cost of ownership.

Keywords: *In-Memory, Columnar database, Business Intelligence, Real time, Performance.*

1 INTRODUCCIÓN. -

Los negocios actuales están propensos a un mercado con cambios extremadamente rápidos lo que demanda cada vez con mayor frecuencia el análisis de datos en tiempo real soportados por sus sistemas OLTP. Los sistemas de Inteligencia de Negocios (BI por sus siglas en Inglés) pueden generar conocimiento basados en esa data a tiempo real, más que los reportes basados en data histórica, lo que permitirá una más rápida y mejor toma de decisiones (Baboo & Kumar, 2013). En el campo de la venta minorista conocida con el anglicismo *retail*, en los años recientes esta industria se ha visto alterada por el empoderamiento del consumidor al usar tecnologías emergentes. Por ejemplo, los clientes usan teléfonos inteligentes para comparar precios mientras están de compras en las tiendas o basan sus decisiones de compra

según retroalimentación de redes sociales, a eso se suma una siempre creciente lista de *retailers on-line* que ofrecen productos con precios competitivos (MacKenzie, Meyer, & Noble, 2013). Las investigaciones de mercado indican que esta tendencia de negocio y de tecnología han cambiado completamente el *retail* tal como lo conocemos ahora e incluso para los próximos años se predicen mayores cambios de los que sucedieron en el siglo pasado (Davison & Burt, 2014). Por ello, la agilidad se torna en un reto particular en el ámbito del BI, considerando que la arquitectura subyacente para la toma de decisiones empresariales se soporta en datawarehouse (DWH) los cuales no se basan en agilidad, sino en confiabilidad y robustez (Knabke & Olbrich, 2015).

Para los sistemas empresariales, los dispositivos y mecanismos de almacenamiento de datos han sido fundamentales ya que en gran medida ellos determinan la rapidez de respuesta a los requerimientos por parte de esas aplicaciones, además que son claves para evitar la pérdida de datos críticos en las operaciones de negocio (Kabakus & Kara, 2016). El tiempo de acceso a disco llegó a ser la razón dominante para un bajo rendimiento en las aplicaciones de base de datos, a menudo provocando que la unidad central de procesamiento (CPU) tenga que esperar para que se completen las operaciones de entrada y salida del disco (Serapiglia, 2010).

Bajo tal entorno, en los últimos años el crecimiento de las bases de datos NoSQL (*Not only SQL*) ha sido notable en buena medida por el hecho de que los tradicionales sistemas de gestión de bases de datos relacionales han tenido dificultades debido al constante crecimiento de la data, las limitaciones de escalabilidad, la pérdida de eficiencia de las consultas (*queries*) sobre grandes volúmenes de datos y la gestión de almacenamiento (Kabakus & Kara, 2016). Una de las características comunes de estas bases de datos es que aprovechan la tecnología In-Memory (IM) que permiten que la data detallada resida completamente en la memoria principal mejorando el rendimiento y la ejecución de operaciones de cálculo (Garg & Majumdar, 2013). A su vez, IM permite dividir los procesos analíticos en piezas manejables a través de computación distribuida y procesamiento paralelo, generando resultados con mayor rapidez a la vez que elimina gran cantidad de preprocesamiento de data con el consiguiente ahorro de recursos (Garber, 2012).

Varios han sido los factores que han contribuido al crecimiento y uso de la tecnología In-Memory (IM), no obstante los de más relevancia residen en la disponibilidad de procesadores de 64 bits, el paralelismo masivo a través de CPUs multinúcleo y la constante reducción de precios de la memoria (Plattner, 2009). Entre los beneficios resultantes más significativos de esta plataforma es el largo espacio de memoria direccionable, cuya capacidad teórica es la siguiente (Baboo & Kumar, 2013):

TABLA 1. MEMORIA DIRECCIONABLE

Sistema	Memoria Direccionable
32-bits	4 GB
64-bits	18 billones GB (18 Exabytes)

Según (Baboo & Kumar, 2013) los límites de memoria direccionable para sistemas de 64 bits están dados por las capacidades de la plataforma y algunos alcances de costo de la misma memoria. Dado que éstos están en constante mejora, la capacidad de los sistemas IM continua en crecimiento y el costo por gigabyte de memoria continua a la baja.

Como consecuencia de todo lo mencionado, el objetivo de este trabajo es analizar nuevos esquemas de bases de datos, que soportadas por tecnologías emergentes como In-Memory permiten la implementación de aplicaciones empresariales con menores recursos, en tiempos más reducidos y con resultados cuya agilidad sobrepasa a los esquemas tradicionales. Para ello, se realizará una revisión de trabajos recientes relacionados con este aspecto y se compararán diferentes perspectivas para identificar aspectos ya consolidados y cuestiones sin resolver en relación con esta tecnología. En la sección 3 se analizará la tecnología IM y los elementos de infraestructura tecnológica que ha propiciado su crecimiento. Se estudiará las bases de datos NoSQL y columnares, y sus características de alto rendimiento y disponibilidad. Además, se revisan cuatro casos de estudio sobre sistemas IM que han tenido un significativo impacto en la agilidad. En la sección 4 se presentan las conclusiones agrupadas en relación con la plataforma, a las bases de datos NoSQL, a las aplicaciones empresariales y el análisis en tiempo real, y en cuanto a los casos estudiados.

2 LA TECNOLOGÍA IM Y SU IMPACTO EN LA AGILIDAD.

Según (Garber, 2012) y (Baboo & Kumar, 2013) la combinación de avances en desarrollo

de hardware y la tecnología IM en aplicaciones analíticas, permitirán el procesamiento de grandes volúmenes de datos en tiempo real para proveer resultados instantáneos en la toma de decisiones. Este rápido acceso y análisis de la data posibilita implementaciones más rápidas y reducción del costo total de propiedad (TCO por sus siglas en Inglés) debido al uso menor de hardware, mantenimiento, pruebas y recursos (Garg & Majumdar, 2013), (Garber, 2012).

Hace algunas décadas las limitaciones computacionales obligaron a que la analítica multidimensional requiera de medidas fijas precalculadas. Para facilitararlo, la respuesta era separar físicamente la base de datos transaccional del DWH y con ello evitar degradaciones de rendimiento. No obstante, los sistemas ahora pueden ser desarrollados de tal forma que permitan consultas en tiempo real hacia la base de datos transaccional en vivo sin pérdida de rendimiento, lo que posibilita el BI en tiempo real sin necesidad de un DWH. Un principio básico en el esquema anterior fue segregar el procesamiento operacional del analítico en capas con almacenes de datos

TABLA 2. TIEMPO TRANSCURRIDO (ms) Y USO DE MEMORIA (MB) EN BASES DE DATO SIN-MEMORY PARA LA OBTENCIÓN DE UN CONUNTO COMPETO DE DATOS.

Base de Datos ¹	Número de registros				Número de registros			
	1,000	10,000	100,000	1,000,000	1,000	10,000	100,000	1,000,000
	Tiempo transcurrido (ms)				Uso de memoria (MB)			
Redis	10	9	11	11	2.1	2.1	2.2	2.2
MongoDB	9	15	9	8	1.3	1.3	1.1	0.8
Cassandra	9	32	24	54	0.6	0.6	1.3	1.3
H2	9	14	18	20	1.2	1.1	9.7	10.6

Fuente: (Kabakus & Kara, 2016)

2.1 LAS BASES DE DATOS NOSQL.

Se debe considerar que en un sistema de gestión de bases de datos relacionales (RDBMS) el rendimiento se ve impactado en la medida que aumenta el tamaño de la data; esto lleva a problemas de concurrencia y causa bloqueos. En contraposición, las bases de datos NoSQL gestionan la data en memoria en lugar de mecanismos no-volátiles como son los discos, lo que permite incrementar significativamente la velocidad de consultas comparado con el acceso lento por medio de I/O a disco (Kabakus & Kara, 2016).

Se han efectuado interesantes experimentos sobre bases de datos NoSQL en memoria para medir el impacto en el rendimiento; (Kabakus & Kara, 2016) han efectuado mediciones de escritura de pares clave-valor, de lectura de una

independientes y principios de diseño radicalmente diferentes. Como resultado era necesario construir y mantener dos sistemas de hardware separados, uno para la base transaccional y otro para analítica de BI, con consecuencias de escenarios más costosos (Serapiglia, 2010).

La tecnología IM elimina el cuello de botella que significa el acceso a disco que es inherente a las bases de datos tradicionales, lo que resulta en mejoras de rendimiento en procesos de analítica. Además, se reduce la dependencia en el almacenamiento de datos en cubos de Procesamiento analítico en línea (OLAP) y en agregaciones, lo que a su vez resulta en implementaciones más rápidas y menor TCO (Baboo & Kumar, 2013).

clave dada, eliminación de un par clave-valor, y de recuperación de todo un conjunto de datos. Estas pruebas incluyeron a una base de datos relacional cuya particularidad es que almacena los datos en memoria en lugar de en disco. Con ello se pudo concluir que el modelo de base de datos tiene un impacto importante en el rendimiento. El último experimento que implica recuperación de todo un conjunto de datos, permite reconocer la rapidez de obtención de todos los datos y el consumo de memoria para hacerlo; los resultados encontrados se listan en la Tabla No.2.

En concordancia con el Teorema de CAP (*Consistency, Availability, Partition Tolerance*) también conocido como Teorema de Brewer (Brewer, 2012), reconoce que en un sistema de computación distribuido es imposible garantizar las tres propiedades simultáneamente y solo dos de las tres son posibles. Los sistemas NoSQL

¹ Redis, MongoDB, Memcached y Cassandra son bases de datos NoSQL. H2 es base relacional

no garantizan ACID (Atomicity, Consistency, Isolation, Durability), contrario a lo que ocurre en las bases de datos relacionales. Brewer señala que, así como las propiedades ACID de los sistemas de bases de datos relacionales proveen consistencia, las propiedades BASE (*Basically Available, Soft-state, Eventually Consistent*) proveen disponibilidad. Conociendo estas restricciones, se sugiere utilizar como criterio de selección los requerimientos que se consideren más críticos para el negocio (Jaramillo Valbuena & Londoño, 2014), (Formia, y otros, 2015). La idea de las bases de datos NoSQL es disminuir algo de consistencia con el objeto de obtener mayor disponibilidad, escalabilidad y alto rendimiento (Kabakus & Kara, 2016).

2.2 TRANSICIÓN EN EL DISEÑO DE BASE DE DATOS. HACIA UN ALMACENAMIENTO COLUMNAR.

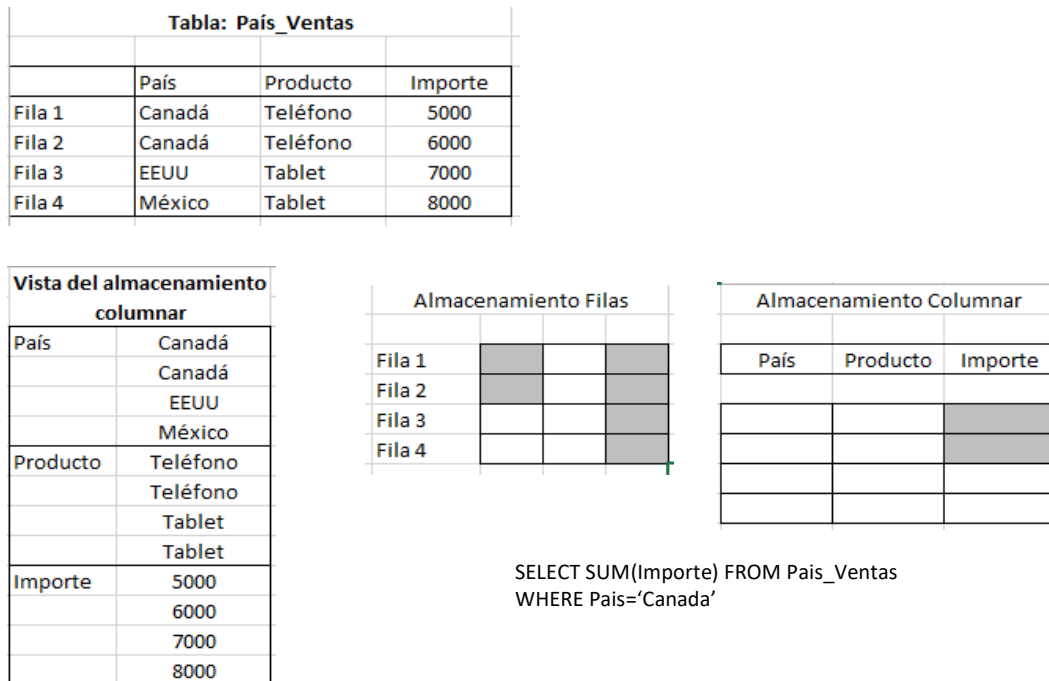
Con el constante crecimiento del volumen de datos en los negocios, la tecnología tradicional ha dividido los sistemas en dos: i) aquellos basados en OLTP cuya base de datos está orientada a estructuras optimizadas de escritura, es decir su diseño se enfoca en gestionar rápidamente grandes cantidades de pequeñas y

simultáneas transacciones; y, ii) aquellos basados en OLAP que proveen de analítica sobre la data almacenada por los sistemas OLTP, éstos están diseñados para gestionar transacciones complejas, con barridos secuenciales que abarcan pocas columnas pero muchas filas de datos.

Los sistemas OLAP requieren data optimizada para lectura lo que ayudará a mejorar el rendimiento en las consultas (Baboo & Kumar, 2013) y (Plattner, 2014) Una forma más eficiente del esquema OLAP es a través de un diseño columnar IM para las tablas ya que usan

técnicas de compresión muy eficientes (tales como prefijos de codificación, codificación de longitud en ejecución, codificación de *cluster*), lo que posibilita realizar barridos muy rápidos de columnas evitando agregaciones y otros diseños complejos en la base de datos (Baboo & Kumar, 2013). Según (Plattner, 2015), las bases de datos columnares

Figura 1. Almacenamiento columnar y eficiencia en operaciones de consulta
Fuente: Adaptado de (Baboo & Kumar, 2013)



En este formato de almacenamiento el contenido de cada columna es almacenado en secuencia de datos

In-Memory rompen la creencia común de que éstas no son adecuadas para la gestión transaccional en aplicaciones empresariales.

Sugiere que este concepto ha sido adoptado por la industria y la academia (Zhang, Chen, Chin Ooi, Tan, & Zhang, 2015), liderando un cambio dramático para la simplificación de los modelos de datos, reducción de la complejidad transaccional, alta flexibilidad en las consultas, y finalmente un impresionante rendimiento a un menor TCO.

En otro artículo (Plattner, 2014) afirma que es factible ejecutar consultas analíticas sobre grandes *sets* de datos directamente sobre esquemas libres de redundancia, eliminando la necesidad de mantener tablas con agregaciones preconstruidas. El hecho de mantener agregaciones preconstruidas tiene tres desventajas a considerar; i) escasez de flexibilidad, lo que no permite reaccionar ante los cambios organizacionales, ii) añade complejidad a los sistemas empresariales, y

iii) incrementa el costo para la inserción de data ya que las vistas agregadas deben mantenerse actualizadas. El almacenamiento columnar IM supera estas desventajas y proporciona ventajas de rendimiento aún sobre configuraciones basadas en filas que se adaptan mejor para rápidas inserciones.

(Zhang, Chen, Chin Ooi, Tan, & Zhang, 2015) indican que a través del cambio de la capa de almacenamiento de datos del disco hacia la memoria principal puede conducir a una mejora teórica de más de 100 veces en términos de tiempo de respuesta y rendimiento (*throughput*). Este cambio incita a repensar el diseño de los sistemas tradicionales, especialmente para bases de datos, en los aspectos de diseño de datos, índices, paralelismo, control de concurrencia, procesamiento de consultas, tolerancia a fallas, etc. La utilización de CPUs modernos y la optimización de la memoria juegan un rol significativo en el diseño de sistemas IM.

Según (Plattner, 2014) el diseño tradicional de bases de datos en filas será reemplazado con una arquitectura común de base de datos que soporte tantas aplicaciones empresariales transaccionales y analíticas. Afirma que las dos decisiones de diseño fundamentales son, i) almacenar la data en formato columnar, y, ii) y mantenerla permanentemente residente en memoria principal. Esto puede ser posible debido al desarrollo de hardware que habilita grandes cantidades de memoria a precios

menores, junto al crecimiento de la computación paralela y el multiprocesamiento; con ello se puede barrer la data en la memoria principal con rendimientos antes impensados. Este desarrollo de hardware en combinación con el formato columnar de tablas y las técnicas de compresión, permiten mejorar la velocidad de barrido, además de reducir el tamaño de la data y dividir la carga entre múltiples núcleos.

Cuando se analizan los dos sistemas, OLTP y OLAP, el escenario en que el primero es más rápido sólo es aquel irreal cuando se pueden anticipar todas las agregaciones y éstas no cambian. En cualquier momento una consulta *ad-hoc* requiere alguna agregación que no ha sido preconstruida, en este caso una arquitectura basada en filas es en órdenes de magnitud más lento. Y el cambio de una estructura agregada podría implicar un esfuerzo de procesamiento con un costo muy alto. Por tanto, una arquitectura columnar puede considerarse como más rápida para el procesamiento de datos cuando se toman en consultas anticipadas y *ad-hoc* (Plattner, 2014).

2.3 CASOS DE ESTUDIO.

A continuación, se estudian algunos casos publicados en los que los sistemas IM han tenido un importante impacto en la agilidad.

El primer caso de estudio se denomina “*An associative engines based approach supporting collaborative analytics in the Internet of cultural things*”. Este estudio ilustra un enfoque integrado que combina *Business Intelligence* (BI), *Big Data* e Internet de las Cosas (IoT) incluyendo contenido estructurado y no estructurado (Chianese, Marulli, Piccialli, Benedusi, & Jung, 2017). El proyecto colecciona datos de una zona turística rica en activos culturales; a través de la tecnología se sigue la pista al comportamiento de los visitantes por medio de sensores, micrófonos, cámaras, identificadores de radio frecuencia (RFIDs), dispositivos móviles y otros. Además, se puede obtener información adicional de un análisis anonimizado del contenido de redes sociales, considerando una fuente casi de tiempo real y que combina otros atributos tales como la georeferencia, la semántica y la temporalidad. El sistema recoge información detallada durante las interacciones con los visitantes y graba los datos abiertos en archivos de log con formato JSON. El proyecto ha implementado una

solución de BI con tecnología IM (Qlik²) en la que se ha alcanzado una alta velocidad de procesamiento sin necesidad de pre-diseñar y pre-agregar datos en un almacenamiento persistente; la exploración de datos se puede realizar en un espacio multidimensional, con operaciones de *roll-up*, *drill-down*, *slicing & dicing*, sin ser obstaculizados por mecanismos jerárquicos.

El segundo se refiere al caso denominado “*In-Memory Technology and the Agility of Business Intelligence – A Case Study at a German Sportswear Company*” (Knabke & Olbrich, 2015). En este caso se analiza si las características de un DWH impactan en la agilidad del BI en el estudio de una empresa manufacturera considerada como una de las grandes industrias de ropa deportiva a nivel mundial; ésta ha implementado un cambio de arquitectura desde una base de datos basada en disco (DRDB) hacia una base de datos In-Memory (IMDB). El escenario inicial consistía en dos principales sistemas BI basados en DWH: un sistema DWH contenía toda la información relacionada con el *retail*; y, el resto de la información relacionada al BI (manufactura, finanzas, etc.) se alojaba en un segundo sistema DWH. Ambos sistemas basados en DRDB contenían componentes adicionales necesarios para alcanzar los requerimientos mínimos de hardware; sin éstos no era posible ejecutar las tareas de reporte y consultas. En el nuevo panorama solamente existía un sistema BI basado en DWH el cual está completamente soportado en IMDB (ver figura 2).

En este estudio (Knabke & Olbrich, 2015) sugieren que el proceso típico inicia con la obtención de la data desde los sistemas fuente hacia la capa de extracción. En el segundo proceso, la data es limpiada, consolidada y armonizada, y después enriquecida con lógica funcional específica. Si un DWH es basado en DRDB, usualmente requiere una capa de reporte por efectos de rendimiento (denominada *data mart*). En relación al rendimiento se ha analizado el proceso de reporte de ventas y márgenes de fin a fin, desde la extracción de los sistemas fuente hasta que la data está disponible para el análisis.

² 1993-2017 QlikTech International AB, All Rights Reserved

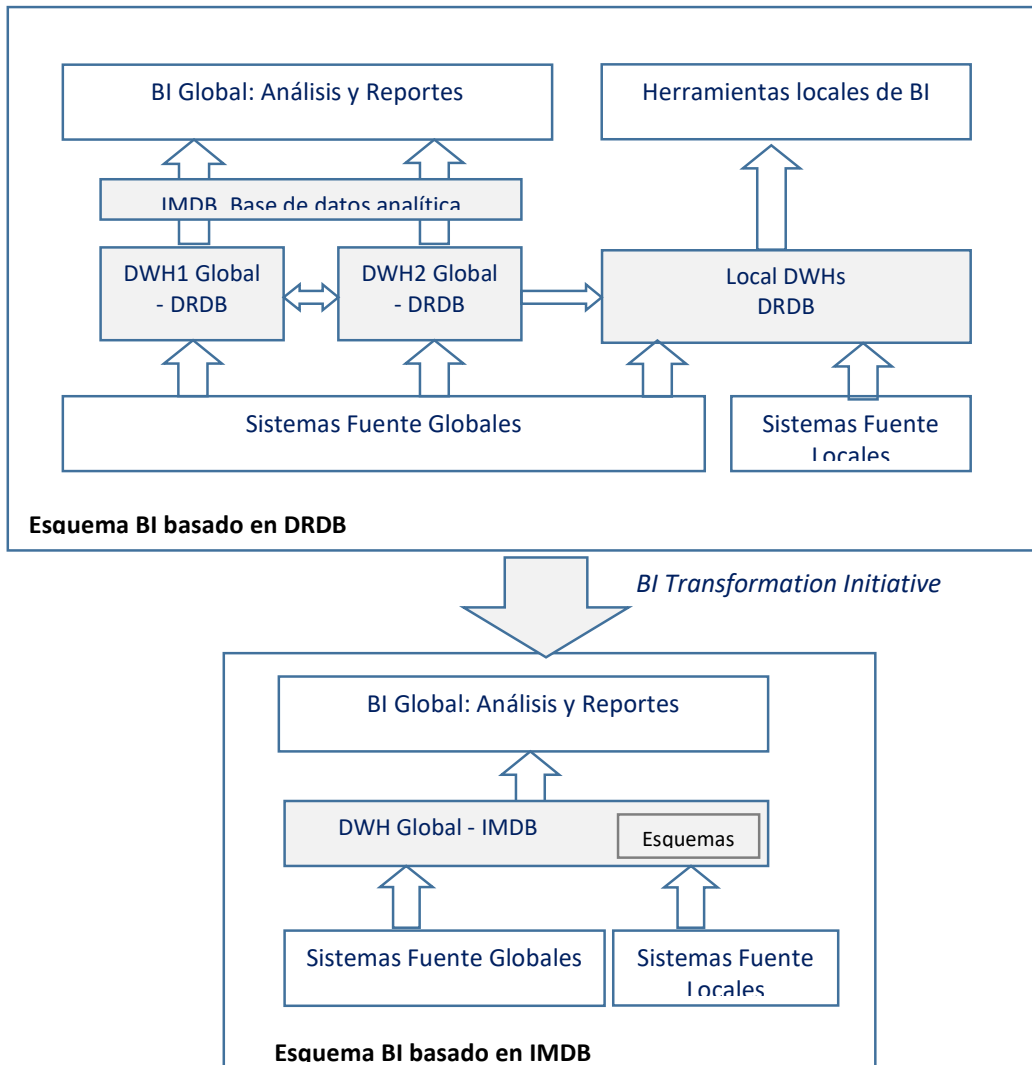


Figura 2. Panorama del Sistema
Fuente: Adaptado de (Knabke & Olbrich, 2015)

En total este proceso decreció de 01:51 horas a 9 minutos (-92%) para un set comparable de datos (aproximadamente 0.2 millones de registros). Además, el sistema DRDB requirió una capa de *data mart* e infraestructura técnica adicional por razones de rendimiento; esta capa no es necesaria en la nueva plataforma (-100%). El resumen se observa en la tabla 3.

Estos resultados obtenidos por (Knabke & Olbrich, 2015) en su estudio indican que en la práctica el seguimiento de todos los criterios de Inmon no resultan en ningún efecto desventajoso para un enfoque de DWH. Con respecto al conflicto del requerimiento de agilidad en el enfoque común del BI basado en DWH, éste parece depender de las variables

asociadas y de la tecnología subyacente. Parece que el concepto básico del DWH tiene un fuerte impacto negativo en la agilidad del BI si está basado en DRDB. Además, la tecnología IM posibilita la implementación de métricas y KPIs que no fueron factibles antes debido a limitaciones de rendimiento. Expertos encuestados en el estudio mencionaron la simplicidad y el reducido tiempo para alcanzar cambios en el sistema BI basado en IMDB como uno de los principales logros de la transformación del BI. La tecnología IMDB subyacente permite cálculos al momento (*on the fly*) y de este modo incrementa el valor al cliente y el beneficio al negocio. Finalmente, los expertos encuestados opinan que esta tecnología

permite un alto nivel de auto-servicio (*self service BI*).

TABLA 3. COMPARACIÓN PROCESO ETL ENTRE EL ANTIGUO Y NUEVO PANORAMA BI.

Proceso / Flujo de datos	Sistema BI basado en DRDB (hh:mm:ss)	Sistema BI basado en IMDB (hh:mm:ss)	Diferencia
Capa Data Mart	01:41:31	00:00:00	-100%
Capa Corporativa y Propagación	00:08:05	00:08:36	+6%
Capa Adquisición	00:01:44	00:00:40	-62%
Total	01:51:19	00:09:17	-92%

Fuente: (Knabke & Olbrich, 2015)

El tercer caso de estudio se denomina “*In-Memory Analytics: Business Intelligence Without a Datawarehouse*” (Serapiglia, 2010). El caso parte del crecimiento tecnológico y la continuidad de la Ley de Moore (Moore, 1965) y cuestiona la necesidad original de separar los requerimientos analíticos del BI del conjunto de datos transaccionales. El caso analiza cómo una organización ha podido alcanzar la unión de estos dos lados en un solo sistema y realmente pueda ser valorable un BI en tiempo real sin un DWH. La organización estudiada opera como recolector y reciclador de neumáticos en los EEUU, con una red de 20 instalaciones productivas, recicla aproximadamente un tercio del total del mercado y provee recogida de producto de costa a costa en más de 60,000 localidades. Esta organización ha visto incrementada su capacidad a través de la adquisición de pequeñas empresas a lo largo del país, las cuales proveyeron de mayor volumen de negocio y clientes, pero de escasa o ninguna infraestructura de tecnologías de información (TI). Sus sistemas operaban en forma local, en cada sitio, y luego se cargaban en un punto central; luego su data era enviada hacia un DWH sobre el que se construían cubos basados en los requerimientos de usuarios. Esto generó con el tiempo que el sistema se vuelva cada vez más lento, con continuas fallas en los procesos *batch*, necesidades de limpieza y re-ejecución, entre otras dificultades. Las alternativas propuestas a estas dificultades consistían en dos opciones: la primera era construir dos sistemas de bases de datos separados con uno dedicado a la recolección de datos y otro a un almacén dedicado a la analítica. La organización decidió la segunda opción que se constituía de un solo sistema más poderoso que fuera capaz de

ejecutar el software empresarial para que ambas tareas (recolección y analítica) puedan ser simultáneas. El sistema costó a la compañía aproximadamente 40% menos que el costo proyectado de la alternativa con múltiples servidores / DWH. A la fecha del estudio se tenía una base de datos centralizada con un peso de 125 GB, la cual residía completamente en memoria RAM. El sistema de reporte descansaba directamente sobre una base en vivo para extraer información bajo demanda y los cálculos se realizan en el momento que se requieren (Serapiglia, 2010).

El cuarto caso analizado se denomina “*A Common Database Approach for OLTP and OLAP Using an In-Memory Column Database*” (Plattner, 2009). Este estudio teórico-práctico analiza la aplicación empresarial SAP instalada en una industria cervecera alemana con más de 5 años de registro de datos. Se realiza una lectura completa (*full scan*) de una tabla de documentos contables con 160 atributos y un aproximado de 34 millones de registros. En una base de datos de filas, un millón de tuplas de esta tabla particular consume alrededor de 1GB de espacio, por lo que la medida total fue de 35 GB. El equivalente almacenamiento columnar de la tabla solo pesaba 8 GB debido a la eficiente compresión vertical entre columnas. En la figura 3 se indica en forma esquemática que el almacenamiento en filas con compresión horizontal no puede competir si el procesamiento es orientado a conjuntos y requiere operaciones de columnas.

El estudio de (Plattner, 2009) sugiere que entre los principales beneficios está el dado por el uso del procesamiento paralelo. Las operaciones de barrido a través de una o más columnas son aprovechadas al máximo, ya que éstas pueden dividirse fácilmente en partes iguales y ser distribuidas en múltiples núcleos. Todos los cálculos a nivel de tuplas automáticamente serán paralelizados, ya que son independientes entre sí. Incluso si solamente unas pocas tuplas equiparan con la selección, no es necesario introducir índices ya que la velocidad de barrido (*scanning*) es muy alta, especialmente si el procesamiento paralelo entre múltiples núcleos está activo. En los actuales CPUs se puede esperar un procesamiento de 1MB per ms y con el procesamiento paralelo en 16 núcleos más de 10 MB per ms. En contexto implica que, con una sola dimensión comprimida en 4 bytes, se puede escanear 2.5 millos de tuplas en un 1ms.

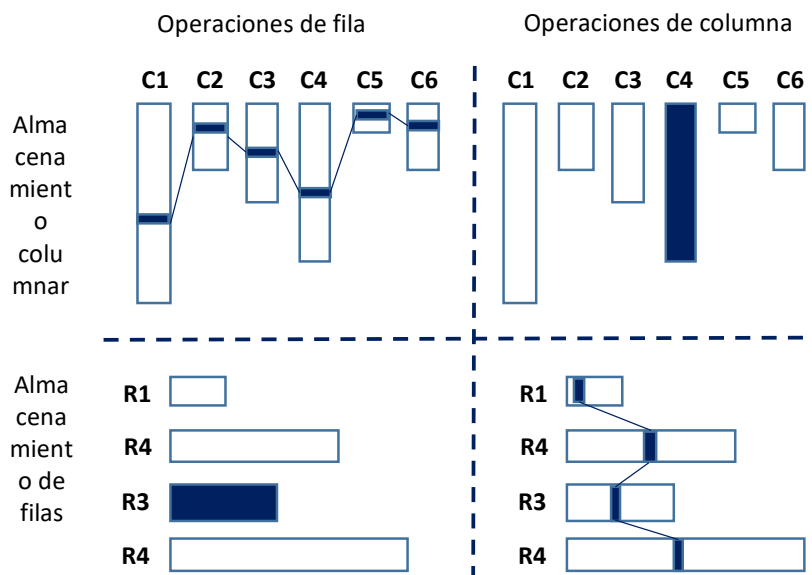


Figura 3. Acceso a datos en almacenamiento en Filas y Columnas
Fuente: Adaptado de (Plattner, 2009)

Otros de los elementos importantes que arroja el estudio de (Plattner, 2009) es el relacionado al consumo de memoria cuando se tiene almacenamiento en filas o columnar de una tabla, donde observa una importante diferencia en la tasa de compresión. Su estudio indica que en varios análisis en data real presentan una típica tasa de compresión de 20 para almacenamiento columnar y 2 para almacenamiento en filas en el disco.

3 CONCLUSIONES.

El estudio descrito permite reconocer algunas conclusiones de interés bajo diferentes aristas, las cuales se exponen y discuten a continuación:

En cuanto a la plataforma:

Las características de rendimiento que ofrece la tecnología IM que subyacen sobre plataformas de 64 bits, multiprocesadores y computación paralela, posibilitan que las organizaciones puedan reducir el TCO y el esfuerzo en mantenimiento, logrando mejores rendimientos en las consultas. Además, posibilita el acceso de esta tecnología a industrias de tamaño medio o pequeño (Baboo & Kumar, 2013).

En cuanto a las bases de datos NoSQL:

Las bases NoSQL que almacenan los datos en memoria se constituyen en elementos críticos para la agilidad en los sistemas de análisis de datos debido a algunas características que las componen tales como el procesamiento de grandes volúmenes con mayor rapidez que los sistemas relacionales por sus arquitecturas altamente escalables, la flexibilidad en las estructuras de datos, y la baja latencia y alto rendimiento (Kabakus & Kara, 2016).

Las bases de datos de tipo SQL y NoSQL ofrecen diferentes características, no obstante, si se permite alguna flexibilidad en la consistencia, las bases NoSQL pueden ofrecer mayor disponibilidad, escalabilidad y alto rendimiento (Kabakus & Kara, 2016).

En cuanto a las aplicaciones empresariales y el análisis en tiempo real:

La tecnología IM es un habilitador de la agilidad en los procesos de analítica al permitir tiempos de respuesta extremadamente rápidos, incluso sobre grandes volúmenes de datos. Esto ha hecho que muchos fabricantes desarrollen

soluciones que ofrecen entre 10 y 100 veces mejor rendimiento que las soluciones analíticas tradicionales sobre discos (Garber, 2012), lo que permite a las organizaciones obtener mayor conocimiento dentro de grandes volúmenes de data transaccional. A su vez, la tendencia de los negocios pone énfasis en el reporte en tiempo real y mayor profundidad en analítica; todo ello habilitará a la organización a tomar decisiones de negocio más efectivas (Baboo & Kumar, 2013).

Estas herramientas basadas en tecnología IM han permitido alcanzar una alta velocidad de procesamiento con ninguna o una mínima necesidad de pre-diseñar y pre-agregar datos en un almacenamiento persistente, permitiendo la exploración de datos en espacios multidimensionales sin ser obstaculizados por estructuras jerárquicas.

En cuanto a los casos estudiados:

El primer caso (Internet de las cosas culturales) permite reconocer que en la actualidad el uso de redes sociales y la información no estructurada proveniente de una variada cantidad de dispositivos han impulsado la necesidad de análisis en tiempo real. Por ello, aquellas soluciones que permitan una mayor agilidad y posibilidad de análisis en tiempo real, serán generadores de ventajas competitivas. Estas soluciones deberán posibilitar un enfoque integrado combinando *Business Intelligence* (BI), *Big Data* e Internet de las Cosas aplicado a recursos de información de fuentes heterogéneas que incluyen data de tipo estructurado y no estructurado. En este sentido, la tecnología IM será un habilitador de agilidad para el análisis de información.

En el segundo caso (Agilidad con IM), los resultados del estudio indican que la tecnología IM aparece como una tecnología habilitadora para un BI ágil. Se reconoce también que algunas características de sistemas BI basados en DWH consiguen efectos positivos en agilidad siempre que subyacen sobre tecnología IM; a su vez el concepto básico del DWH tiene un fuerte impacto negativo en la agilidad cuando está basado en DRDB. Esta característica de agilidad se observa en el resultado del caso de estudio al disminuir el tiempo total de procesamiento en un 92% cuando se compara un escenario de DWH sobre DRDB e IMDB, además de eliminar la necesidad de infraestructura adicional para la capa de reporte (Knabke & Olbrich, 2015).

La tecnología IM posibilita la implementación de métricas y KPIs que no fueron factibles antes debido a limitaciones de rendimiento. Además, posibilita cálculos al momento (*on the fly*) y combinados con el poder computacional permiten generar un mayor beneficio al usuario y al negocio, propiciando incluso un alto nivel de auto-servicio y análisis *ad-hoc*.

En el tercer caso de estudio (Agilidad con IM sin DWH), se observa que para compañías medianas y pequeñas se pueden encontrar eficiencias operacionales si son soportadas por tecnología IM. A través del análisis en tiempo real facilitado por la tecnología IM ejecutada sobre una base transaccional en vivo, las compañías podrían ser capaces de tomar decisiones de negocio eficientes con rapidez y exactitud. El caso muestra que una arquitectura de este tipo puede generar un menor TCO (-40% en este estudio) ya que requiere menor infraestructura (Serapiglia, 2010).

A su vez, tal arquitectura constituye una ventaja sobre los sistemas tradicionales OLAP que requieren pre-agregaciones y no prestan facilidad para añadir métricas fuera del cubo original. En cambio, IM ha permitido el análisis de datos en tiempo real sobre una base de datos transaccional en vivo, sin pérdida de rendimiento; en consecuencia, ha otorgado mejor soporte a las decisiones de negocio con eficiencia, rapidez y exactitud. Esto posibilita el BI en tiempo real sin la necesidad de un DWH (Serapiglia, 2010).

En el cuarto caso (BD columnares e integración OLAP y OLTP), el paradigma de la separación de la data empresarial en diferentes capas físicas, una para OLTP y otra para OLAP fue necesaria para permitir un adecuado rendimiento de las aplicaciones; no obstante, dado el crecimiento del procesamiento paralelo y la disponibilidad de memoria RAM, ambos a precios cada vez menores, podrían dejar obsoleto este enfoque. Ello implicaría un solo sistema tanto para las operaciones transaccionales como para las analíticas, generándose los cálculos sobre demanda y a una muy alta velocidad (Plattner, 2009).

El almacenamiento columnar permite una alta compresión de los datos y ésta se ve favorecida especialmente por el procesamiento paralelo. Las operaciones de escaneo tienen un diferencial de velocidad en orden de magnitud comparadas con el almacenamiento en filas, a tal punto que incluso el uso de índices podría quedar obsoleto; además estas operaciones pueden dividirse

fácilmente en partes iguales y ser distribuidas en múltiples núcleos. A esto se añade que el consumo de memoria en una estructura de almacenamiento columnar es mucho más eficiente llegando a una tasa de compresión de 20 versus 2 comparado con la tasa en almacenamiento en filas (Plattner, 2009).

La disponibilidad de grandes capacidades de memoria principal ha sido una de las tendencias que hacen que la propuesta de una arquitectura de base de datos columnar IM sea una solución viable. En el futuro, la persistencia primaria podría ser almacenada en memoria no-volátil para reducir significativamente los tiempos de recuperación con mecanismos seguros y consistentes (Plattner, 2014).

4 REFERENCIAS.

- Baboo, S., & Kumar, R. (Mayo de 2013). Next Generation Data Warehouse and In-Memory Analytics. *International Journal of Computer Applications*, 69(18), 25-30. Recuperado el 16 de Agosto de 2017, de <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.403.9732&rep=rep1&type=pdf>
- Brewer, E. (2012). CAP twelve years later: How the "rules" have changed. *Computer*, 45(2), 23-29.
- Chianese, A., Marulli, F., Piccialli, F., Benedusi, P., & Jung, J. E. (2017). An associative engines based approach supporting collaborative analytics in the Internet of cultural things. *Future Generation Computer System*(66), 187-198. doi:<https://doi.org/10.1016/j.future.2016.04.015>
- Davison, J., & Burt, M. (2014). *Top Retail Business and Technology Trends*. Recuperado el 16 de Marzo de 2015, de Gartner: <https://www.gartner.com/doc/2727618/top-retail-business-technology-trends>
- Formia, S., Vivas, L., Cambarieri, M., García Martínez, N., Muñoz, H., & Petroff, M. (2015). *Tratamiento de Grandes Volúmenes de Datos en Ciudades Inteligentes*. Universidad Nacional de Río Negro, Laboratorio de Informática Aplicada, Argentina. Recuperado el 22 de Agosto de 2017, de <http://repositorioinstitucional.lia.unrn.edu.ar:8080/jspui/handle/20.500.12049/150>
- Garber, L. (2012). Using In-Memory Analytics to Quickly Crunch Big Data. *Computer*, 45(10), 16-18. Recuperado el 08 de Mayo de 2017, de <http://ieeexplore.ieee.org/abstract/document/6329877/>
- Garg, M., & Majumdar, R. (2013). Turn Hours into Seconds: BI Paradigm with In-Memory Analytics. (A. S. Technology, Ed.) *International Journal of Software and Web Sciences (IJSWS)*, 13(345), 24-27. doi:ISSN 2279-0071
- Jaramillo Valbuena, S., & Londoño, J. M. (2014). Sistemas para almacenar grandes volúmenes de datos. *Gerencia Tecnológica Informática*, 13(37), 17-28. doi:ISSN 1657-8236
- Kabakus, A., & Kara, R. (2016). A performance evaluation of in-memory databases. *Journal of King Saud University - Computer and Information Science*. doi:<http://dx.doi.org/10.1016/j.jksuci.2016.06.007>
- Knabke, T., & Olbrich, S. (2015). In-Memory Technology and the Agility of Business Intelligence – A Case Study at a German Sportswear Company. *Conference: Pacific Asia Conference on Information Systems (PACIS), At Singapore* (págs. 1-18). Singapore: ResearchGate. Recuperado el 27 de Mayo de 2017, de <https://www.researchgate.net/publication/283794267>
- MacKenzie, I., Meyer, C., & Noble, S. (2013). *How retailers can keep up with consumers*. (M. & Company, Editor) Recuperado el 03 de Marzo de 2015, de Mckinsey & Company: http://www.mckinsey.com/insights/consumer_and_retail/how_retailers_can_keep_up_with_cons
- Plattner, H. (2009) 'A common database approach for OLTP and OLAP using an in-memory column database', Proceedings of the 35th SIGMOD international conference on Management of data - SIGMOD '09, pp. 1-2. doi: 10.1145/1559845.1559846
- Plattner, H. (2014) 'The Impact of Columnar In-Memory Databases on Enterprise Systems', in Very Large Data Bases. Hangzhou, China, pp. 1722-1729. doi: 10.14778/2733004.2733074
- Plattner, H. (2015) 'A new architecture for enterprise application software based on in-memory databases.', *Informatik*, 31(1), pp. 1-5. doi: ISBN 978-3-88579-640-4
- Serapiglia, A. (2010). In Memory Analytics: Business Intelligence Without a Datawarehouse. *Issues in Information Systems*, XI(1), 445-450. Recuperado el 10 de Mayo de 2017, de http://iacis.org/iis/2010/445-450_LV2010_1390.pdf
- Zhang, H., Chen, G., Chin Ooi, B., Tan, K.-L., & Zhang, M. (Julio de 2015). In-Memory Big Data Management and Processing: A Survey. *IEEE Transactions On Knowledge and Data Engineering*, 27(7), 1920-1942. doi:10.1109/TKDE.2015.2427795

Revisión sistemática de literatura: Análisis de riesgos utilizando Redes Bayesianas

Systematic literature review: Risk analysis using Bayesian Networks

Carlos Patricio Pereira Paredes, Oscar Cumbicus-Pineda

Carrera de Ingeniería en Sistemas, Universidad Nacional de Loja - Ecuador, carlos.pereira@unl.edu.ec
Carrera de Ingeniería en Sistemas, Universidad Nacional de Loja – Ecuador, oscar.cumbicus@unl.edu.ec

Resumen: En los últimos años las redes Bayesianas han sido de mucha importancia ya que es la herramienta principal para la estimación en análisis de riesgos, pues al usar redes bayesianas se obtiene una representación gráfica, modelado cualitativo y cuantitativo, inferencia bidireccional, análisis de sensibilidad, incertidumbre, y valores de confianza, lo cual permite una visión amplia, y al existir multiplicidad de información proporciona de manera eficaz los datos necesarios en la toma de decisiones.

El presente trabajo es una revisión sistemática de literatura (SRL), basada en el procedimiento de Kitchenham[1]. Los artículos analizados van desde 2010 a 2017; la búsqueda inicial dio como resultado 1.854 artículos de los cuales se suprimió los que no tenían información relevante para el estudio, obteniendo un total de 154 artículos de revisión, de los cuales se seleccionó 20 para ser analizados.

En la presente investigación se pudo determinar que la inferencia bayesiana es usada con mayor frecuencia para el análisis de riesgos, al igual que el manto o mapeo de Markov, seguido del clasificador bayesiano.

Palabras clave: *bayesian networks, bayes theorem, risk analysis, bayesian inference, systematic review of the literature.*

Abstract. In recent years Bayesian networks are of great importance since it is the main tool for estimating risk analysis, since using Bayesian networks gives a graphical representation, qualitative and quantitative modeling, bidirectional inference, sensitivity analysis, uncertainty, and confidence values.

The present work is a systematic review of the literature of several scientific articles found in the databases IEEE Xplore, ACM Digital Scopus, Dialnet, Science Direct, for this we have based on the procedure of Kitchenham[1]. The articles analyzed ranged from 2010 to 2017, where 1,854 articles were found. Some of the articles that were not included in the study were deleted, and finally 154 articles were reviewed, of which 20 were selected individually.

In the present investigation, it was possible to determine that in the majority of articles studied they are based on other processes so that the decision making in the risk analysis is the best and that of lesser loss of social-economy.

Keywords: *bayesian networks, bayes theorem, risk analysis, bayesian inference, systematic review of the literature.*

I. INTRODUCCIÓN

El análisis de riesgos es tratado como un proceso de optimización, el cual recomienda la mejor alternativa que optimice el interés esperado dadas las observaciones de los factores externos y preferencias. De acuerdo a Thomas[2] existen proyectos que permiten medir el riesgo por medio de la incertidumbre de las relaciones de dependencia existentes entre las diferentes variables presentes en un sistema, ejemplo de ello es consultar la posible colisión de buques usando conocimientos de expertos y conjuntos de datos de accidentes previos. En la actualidad con los avances de la tecnología, las empresas comerciales, de servicios de

producción, entre otras se ven frente a nuevos retos para intentar tareas cada vez más complejas las mismas que presentan riesgos que además de amenazar el desempeño de sus funciones, pueden ser difíciles de controlar.

Según Lozano[3] la percepción del riesgo es asimilado de distinta manera por cada parte involucrada con la empresa, ya sea interna o externa a la misma y en función de la posición que tenga con la empresa.

En gran mayoría la gerencia es la responsable de la seguridad y del manejo de riesgos, es la encargada de dar respuesta a los intereses de todas las partes, por consiguiente, es la más interesada en la identificación, análisis y reducción de los riesgos.

Y dada esta necesidad se busca poner a disposición de las empresas una herramienta que les apoye en el proceso de toma de decisiones, la cual debe aportar a dar solución a las problemáticas mirando más allá de los aspectos de la seguridad sea esta social, humana, tecnológica, económica, entre otras.

El análisis de riesgos se aplica en las empresas en diferentes casos estos pueden ser: riesgos derivados de las actuaciones de las autoridades administrativas (interpretación de regulaciones gubernamentales, concesión de permisos y licencias, etc.), riesgos técnicos derivados de avances o innovaciones tecnológicas no controladas adecuadamente, riesgos contractuales derivados de mala interpretación de normas, reglamentos y contratos, riesgos financieros dados por cambios de divisas, intereses de capital, mal manejo de fondos, riesgos de impacto socio, humano o ambiental, riesgos de estudio de mercado y competencia.

Una vez descritos los riesgos y conocido sus posibles efectos o daños debe realizarse la gestión, que es equivalente a determinar la mejor alternativa a cada riesgo. La respuesta a los riesgos debe estar en armonía con el resultado económico que puede dar lugar a cada uno. Los riesgos, una vez reconocidos, pueden tener una respuesta inmediata (eliminar, transferir, asegurar, reconocer, etc.), el conocimiento de los riesgos permitirá estar alerta y elaborar oportunamente un plan de contingencia para convertir la situación de riesgo en beneficios para la empresa.

Para la optimización de dichos criterios, en algunas áreas se ha priorizado el uso de redes bayesianas debido a que son un área creciente e importante para la investigación y aplicación en algunos campos de la Inteligencia Artificial[3].

Las redes bayesianas son modelos probabilísticos que relacionan un conjunto de variables aleatorias mediante grafos, además son redes graficas en las que se representan variables aleatorias y las relaciones existentes entre cada una de ellas, que permiten conseguir soluciones a problemas de decisión en casos de incertidumbre.

En el proceso de construcción de una red bayesiana para el análisis de riesgos, es importante definir el problema, teniendo claras todas las características del conflicto, tal que, posteriormente sea posible recolectar los datos que serán las variables fundamentales en el desarrollo del proyecto y se plantee un modelo para obtener la mejor solución que se acople a las necesidades del proyecto [4].

De acuerdo a Rodríguez [5] las ventajas de usar redes bayesianas para el análisis de riesgos son: representación gráfica, modelado cualitativo y cuantitativo, inferencia bidireccional, análisis de sensibilidad, incertidumbre y valores de confianza, a pesar de estas ventajas en proyectos estudiados se ha encontrado algunas limitaciones como: estructura de modelo de dominio, variables ocultas, probabilidades inconsistentes.

Las redes bayesianas tienen una gran capacidad de cuantificar el riesgo y gran flexibilidad para adaptarse a poca información, de esta manera se constituye como una herramienta eficaz para el análisis de riesgos[6]; debido a esto han sido usadas en varias áreas del conocimiento, dentro de las áreas en las que se aplicado las redes bayesianas para determinar el riesgo podemos encontrar las siguientes: medicina[7], seguridad[8], distribución de energía[9], enseñanza asistida por computador[10], telecomunicaciones[11], minerías de carbón[12], seguridad aeronáutica[13], hidroeléctricas[14], mantenimiento predictivo de sistemas de producción eléctrica[15], reducción de tasas de mortalidad causadas por enfermedades[16].

En los sistemas desarrollados en cada una de las áreas descritas en el párrafo anterior se ha podido evidenciar mejoras al reducir costos y tiempos, al igual que ha ayudado a determinar enfermedades con escasa sintomatología como algunos tipos de diabetes y enfermedades cardiopatías, en el trabajo de Marshall[16] se puede ver que la tasa de mortalidad se redujo de 95% a 50%, se puede también apreciar que sirve para determinar políticos corruptos como se muestra en el trabajo de Carvalho[17], donde se predice la corruptibilidad en un 90%.

Tabla 2. Plantilla de Kitchenham, Biolchini para una Revisión Sistemática [1]

II. ALCANCE

El presente trabajo es un revisión sistemática de literatura que examinará el uso de redes bayesianas en el análisis de riegos; se procura analizar este proceso para establecer los métodos bayesianos más recomendados dentro de este campo.

III. METODOLOGÍA

La metodología principal utilizada para realizar la SRL fue el la de Barbara Kitchenham[1], ya que es un medio para evaluar e interpretar de las investigaciones disponibles los estudios más importantes y destacados (llamados estudios primarios o estudios individuales), que van a dar respuesta a una pregunta en particular de la investigación, área temática, o fenómeno de interés; además tienen como objetivo presentar una evaluación razonable de un tema de investigación mediante el uso de una metodología fiable, rigurosa y auditable. [18]

En la Tabla 1, podemos ver un resumen con las principales etapas y actividades de la SRL:

Tabla 1. Etapas de una revisión sistemática[1]

Etapa 1	Planificación de la revisión
	Identificación de la necesidad de la revisión
	Desarrollo de un protocolo de revisión
Etapa 2	Desarrollo de la revisión
	Identificación de la investigación
	Selección de los estudios primarios
	Evaluación de la calidad del estudio
	Extracción y monitoreo de datos
	Síntesis de datos
Etapa 3	Publicación de los resultados

En relación con el método propuesto por Kitchenham[1] y Biolchini[18] se definió una plantilla para el protocolo que nos sirve como guía en las etapas de planificación y ejecución de la SRL.

Dicha plantilla ha sido utilizada como guía para el desarrollo de esta SRL y está organizada en base a una serie de secciones y subsecciones detalladas en la Tabla 2.

Planificación de la revisión
Formulación de la pregunta
<ul style="list-style-type: none"> Foco de la pregunta Amplitud y calidad de la pregunta
Selección de fuentes
<ul style="list-style-type: none"> Definición del criterio de selección de fuentes Lenguaje de estudio Identificación de fuentes Selección de fuentes después de la evaluación Comprobación de las fuentes
Selección de los estudios
<ul style="list-style-type: none"> Definición del criterio de inclusión y exclusión de estudios Definición de tipos de estudio Procedimiento para la selección de los estudios
Ejecución de la revisión
Ejecución de la selección en la fuente "x"
<ul style="list-style-type: none"> Selección de estudios iniciales Evaluación de la calidad de los estudios Revisión de la selección Extracción de información

En el presente trabajo siguió cada uno de los conceptos antes expuestos, se llegó al análisis de los resultados obtenidos y a las conclusiones que nos da la respuesta a las preguntas planteadas en la Tabla 3.

IV. RESULTADOS

A. PLANIFICACIÓN DE LA REVISIÓN

En la planificación se identificó la necesidad de realizar la RSL, se definió los objetivos a alcanzar, se identificó las fuentes que nos proporcionaron los estudios primarios y los criterios de búsqueda que se implementaron, mecanismos de razonamientos utilizados para la inclusión o exclusión de los artículos a ser estudiados y los procesos que nos guiaron para obtener las conclusiones y los resultados.

B. FORMULACIÓN DE LAS PREGUNTAS

DE INVESTIGACIÓN

Esta etapa comenzó formulando las preguntas de investigación en las cuales se centró la presente revisión sistemática, estas se sintetizaron tomando en cuenta el área de interés de investigación y de esta manera se definió el problema a tratar y sus principales características, dichas preguntas se presentan en la Tabla 3.

Tabla 3. Preguntas de Investigación

Preguntas de Investigación	
P 1	¿En qué áreas son usadas las redes bayesianas para el análisis de riesgos?
P 2	¿En el análisis de riesgos, la utilización de redes bayesianas optimiza los procesos y resultados?
P 3	¿Qué métodos bayesianos se usan para el Análisis de riesgos y de qué forma han sido aplicados en los casos de éxito más relevantes?

C. SELECCIÓN DE FUENTES

En este punto se analizó las fuentes que se usaron para realizar la ejecución de la revisión y que son IEEE Xplore, Scopus, ACM Digital, Dialnet y Science Direct. Posteriormente se utilizaron los elementos definidos en la planificación para aplicar el procedimiento de obtención de estudios primarios en cada una de las fuentes seleccionadas; las cadenas de búsqueda que se usaron en esta investigación se presentan en la Tabla 4. **¡Error! No se encuentra el origen de la referencia.**

Tabla 4. Cadenas de Búsqueda

Base de Datos	Identificación	Cadena de Búsqueda
ACM Digital	CB01	(+ bayesian + networks + risk + analysis)
ACM Digital	CB02	(+ bayesian + networks + operational + risk)
ACM Digital	CB03	(+ bayesian + networks + risk + prediction)
IEEE Xplore	CB04	((“Abstract”: bayesian networks) AND (“Abstract”: risk analysis))
IEEE Xplore	CB05	((“Abstract”: bayesian networks) OR (“Abstract”: risk analysis))
IEEE Xplore	CB06	((“Abstract”: bayesian networks) SAME (“Abstract”: risk analysis))
IEEE Xplore	CB07	((“Abstract”: bayesian networks) WHIT (“Abstract”: risk analysis))
IEEE Xplore	CB08	((“Abstract”: bayesian networks) AND (“Abstract”: operational risk))
IEEE Xplore	CB09	((“Abstract”: bayesian networks) OR (“Abstract”: operational risk))
IEEE Xplore	CB10	((“Abstract”: bayesian networks) SAME (“Abstract”: operational risk))
IEEE Xplore	CB11	((“Abstract”: bayesian networks) WHIT (“Abstract”: operational risk))
IEEE Xplore	CB12	((“Abstract”: bayesian networks) AND (“Abstract”: risk prediction))

IEEE Xplore	CB13	((“Abstract”: bayesian networks) OR (“Abstract”: risk prediction))
IEEE Xplore	CB14	((“Abstract”: bayesian networks) SAME (“Abstract”: risk prediction))
IEEE Xplore	CB15	((“Abstract”: bayesian networks) WHIT (“Abstract”: risk prediction))
SCOPUS	CB16	(TITLE-ABS-KEY (bayesian networks) AND TITLE-ABS-KEY (risk analysis)) AND PUBYEAR > 2012 AND PUBYEAR < 2017 AND (LIMIT-TO (SUBJAREA, “COMP”))
SCOPUS	CB17	(TITLE-ABS-KEY (bayesian networks) AND TITLE-ABS-KEY (operational risk)) AND PUBYEAR > 2012 AND PUBYEAR < 2017 AND (LIMIT-TO (SUBJAREA, “COMP”))
SCOPUS	CB18	(TITLE-ABS-KEY (bayesian networks) AND TITLE-ABS-KEY (risk prediction)) AND PUBYEAR > 2012 AND PUBYEAR < 2017 AND (LIMIT-TO (SUBJAREA, “COMP”))
SCOPUS	CB19	(TITLE-ABS-KEY (bayesian networks) OR TITLE-ABS-KEY (risk analysis)) AND PUBYEAR > 2012 AND PUBYEAR < 2017 AND (LIMIT-TO (SUBJAREA, “COMP”))
SCOPUS	CB20	(TITLE-ABS-KEY (bayesian networks) OR TITLE-ABS-KEY (operational risk)) AND PUBYEAR > 2012 AND PUBYEAR < 2017 AND (LIMIT-TO (SUBJAREA, “COMP”))
SCOPUS	CB21	(TITLE-ABS-KEY (bayesian networks) OR TITLE-ABS-KEY (risk prediction)) AND PUBYEAR > 2012 AND PUBYEAR < 2017 AND (LIMIT-TO (SUBJAREA, “COMP”))
Dialnet	CB22	“bayesian networks” AND “risk analysis”
Dialnet	CB23	“bayesian networks” AND “operational risk”
Dialnet	CB24	“bayesian networks” AND “risk prediction”
Dialnet	CB25	“bayesian networks” OR “risk analysis”
Dialnet	CB26	“bayesian networks” OR “operational risk”
Dialnet	CB27	“bayesian networks” OR “risk prediction”
ScienceDirect	CB28	“Abstract”, “Title”, “Keywords”, bayesian networks AND risk analysis”
ScienceDirect	CB29	“Abstract”, “Title”, “Keywords”, bayesian networks AND operational risk”
ScienceDirect	CB30	“Abstract”, “Title”, “Keywords”, bayesian networks AND risk prediction”
ScienceDirect	CB31	“Abstract”, “Title”, “Keywords”, bayesian networks OR risk analysis”
ScienceDirect	CB32	“Abstract”, “Title”, “Keywords”, bayesian networks OR operational risk”
ScienceDirect	CB33	“Abstract”, “Title”, “Keywords”, bayesian networks OR risk prediction”

D. SELECCIÓN DE LOS ESTUDIOS

Una vez definidas las fuentes, fue necesario describir el proceso y el criterio a seguir en la ejecución de la revisión para la selección y evaluación de los estudios primarios. Para ello se definió el proceso completo de selección, así como los criterios de inclusión y exclusión que se va a utilizar y que son los siguientes:

1. CRITERIOS DE INCLUSIÓN

- Artículos científicos publicados a partir del año 2010 hasta el año 2017.
- Artículo científico que en el resumen contenga las palabras claves.
- Artículos científicos que sus títulos tengan relación con el tema de investigación.
- Artículos científicos que sean del área de las ciencias de la computación
- Artículos que hayan sido citados mínimo 2 veces.

2. CRITERIOS DE EXCLUSIÓN

- Publicaciones que en la introducción o resumen no presentaba en un solo párrafo las cadenas de búsqueda.
- Artículos que no contaban con el Identificador digital DOI.
- Artículos duplicados.
- Todas las que no cumplen con los criterios de inclusión.

E. ANÁLISIS DE RESULTADOS

En esta etapa del proceso de la revisión se inició con la clasificación de los resultados obtenidos, luego se realizó una comparación del tipo formal de los principales artículos que se obtuvo, como finalización de este proceso se

definió las conclusiones obtenidas de esta comparación.

F. EJECUCIÓN DE LA RSL

Para empezar esta etapa se muestra un resumen de los estudios que se analizaron en la Tabla 5, la cual define aquellos estudios que se han considerado como relevantes y los que se han tomado en cuenta como estudios primarios. El proceso para la selección de los estudios se describe en la Figura 1, proceso que inicia con la ejecución de la consulta en la fuente de datos seleccionada previamente, de este paso se obtiene un conjunto de estudios a los que se les debe aplicar los criterios de inclusión para llegar a definir los estudios relevantes y luego al aplicar los criterios de exclusión poder definir los estudios primarios.

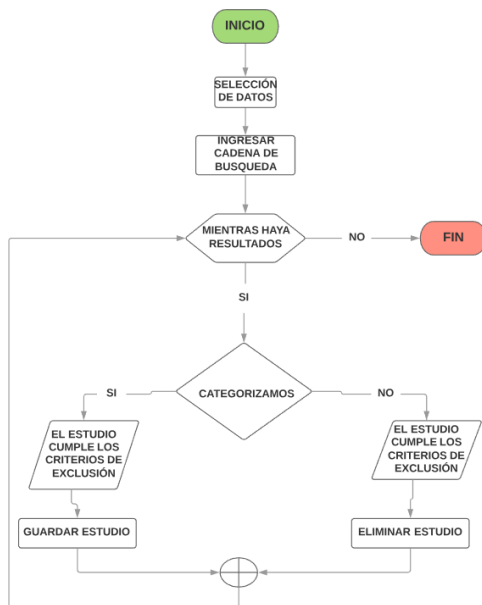


Fig. 1 Proceso de selección de estudios incluidos y excluidos.

Una vez determinados los estudios primarios se realiza una última fase de refinado, la que nos sirve para definir de los estudios importantes, aquellos que se pueden añadir como primarios debido a su contenido y la relación que presentan con los estudios ya seleccionados. Se puede observar que en la presente RSL se han tomado en cuenta para el análisis 154 estudios, de los cuales se consideraron como primarios 20, detalle que se observa en la Tabla 5.

Tabla 5. Artículos analizados

Fuentes	Estudios	Relevantes	Primarios
ACM Digital	138	27	4
IEEE Xplore	1399	71	10
Scopus	187	32	6
Dialnet	120	9	0
ScienceDirect	120	15	0
TOTAL	1854	154	20

V. PRESENTACIÓN DE RESULTADOS

La presentación de los resultados que se obtuvieron durante este proceso se muestra según la clasificación de estudios y la aportación de cada uno de estos, los estudios que se tomaron en cuenta principalmente son:

- Estudios realizados a proyectos en los que se aplicaron Redes Bayesianas para el análisis de riesgos.
- Análisis de los resultados obtenidos en la Aplicación de Análisis de riesgo Utilizando redes Bayesianas

A. DESCRIPCIÓN DEL MARCO DE COMPARACIÓN FORMAL

Existen varios métodos para medir, comparar y evaluar los resultados, según la aplicación que se le va a dar, por lo que hemos elegido la comparación sistemática a destajo, la cual nos permite realizar un sondeo en base a la importancia de cada uno de estos y tomando en cuenta las principales similitudes que muestran los diferentes estudios.

Aunque varios de estos estudios no fueron desarrollados para cumplir el mismo objetivo, ni sobre el mismo dominio específico, es interesante poder analizarlas de algún modo para poder definir qué parte de los procesos de desarrollo pueden reutilizarse en futuras aplicaciones, y exponer cuales de éstas parecen más completas dentro del área del análisis de riesgos utilizando redes bayesianas. La comparación ha sido realizada sobre aquellas propuestas cuyos desarrollos y procesos se muestran bajo un sistema similar o han sido proporcionadas directamente por los autores.



Fig. 2 Estadísticas por año Merezco

En la Fig. 2 se muestra los años en los que se han desarrollado los estudios analizados, lo que nos permite ver la evolución de desarrollo ya que en el 2010 no se muestra mayor cantidad de estudios realizados sobre la aplicación de algoritmos, 2016 es el año en el cual existen más artículos relevantes para nuestra investigación lo que nos indica que los investigadores han dado más prioridad a la investigación para dar solución a los problemas de optimización en el desarrollo de software y así tener productos de calidad y confiables en un menor tiempo. En el año 2017 se muestra un porcentaje de estudios muy bajo ya que algunos artículos demoran la publicación, pero las aplicaciones en las que se realizan los estudios son de mayor profundidad ya que muchos de ellos, toman como base estudios de años anteriores y por tanto no se repiten o se centran en estudios aún no realizados.

C. ESTADÍSTICAS POR IMPACTO

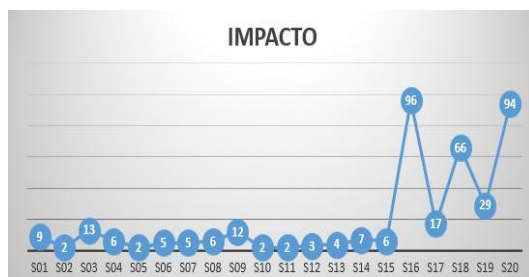


Fig. 3 Estadísticas por Impacto

B. ESTADÍSTICAS POR AÑO

En la Fig. 3, se muestra la incidencia que han tenido los estudios que hemos analizado en otros estudios realizados posteriormente, en dicho gráfico se visualiza que S20 de [20] Yu y S16 de [21] Hossain han mantenido una influencia muy destacada ya que se ha tomado como referencia directa para 94 y 96 estudios respectivamente, los S18 no se queda atrás con un impacto de 66 citas en otros artículos, S19 tiene un impacto de 29 artículos, luego tenemos el S17 con 17 citas y a continuación se encuentra el S03 y S09 con 13 y 12 impactos para otros artículo, en otros artículos la incidencia es más baja, lo cual nos indica que la redes bayesianas tienen un gran impacto para el análisis de riesgos y toma de decisiones

D. ESTADÍSTICA DE REDES BAYESIANAS USADAS EN LOS MODELOS PARA EL ANÁLISIS DE RIESGO

En la Tabla 6 y Fig. 4 se puede observar que el método que se ha utilizado con mayor frecuencia para el desarrollo de los artículos es la Inferencia Bayesiana, en sus dos clasificaciones por propagación de árboles de decisión y por redes multiconectadas los diagramas de influencia, el que le sigue es el Mapeo de Markov seguido del clasificador bayesiano.

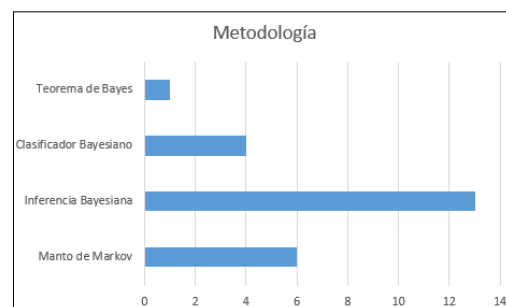


Tabla 6. Modelos Bayesianos Aplicados

N ^o	Nombre del Artículo	Metodología			
		Manto de Mapeo de Markov	Inferencia Bayesiana	Clasificado de Decisiones	Teorema de Bayes
S01	Bayesian risk estimation model for the case of a collision/grounding			X	
S02	Portable Decision Support System for Heart Failure Detection and Medical Diagnosis		X		
S03	Dynamic Bayesian Networks for Situational Awareness in the Presence of Noisy Data		X		
S04	Predicting Cyber Attacks With Bayesian Networks Using Unconventional Signals			X	
S05	Risk Analysis in Cyber Situation Awareness Using Bayesian Approach	X			
S06	A Bayesian-Based Approach for the Short-Term Forecasting of Electrical Loads in Smart Grids. Parte I y Parte II	X	X		
S07	Privacy-Preserving Energy Flowcontrol In Mart Grids		X		
S08	Continuous Dynamic Bayesian Networks for Predicting Survival of Ischaemic Heart Disease Patients	X			X
S09	A Fuzzy Dynamic Bayesian Network-Based Situation Assessment Approach		X		
S10	Modelling Project Complexity driven Risk Paths in New Product Development	X	X		
S11	Optimization of C5.0 Classifier using Bayesian Theory		X		
S12	Using Political Party Affiliation Data to Measure Civil Servants Risk of Corruption		X		
S13	Bayesian Network Application for the Risk Assessment of Existing Energy Production Units		X		
S14	Development of Control Systems for Safety Instrumented Systems	X			
S15	A novel Bayesian hierarchical model for road safety hotspot prediction		X	X	
S16	A Bayesian network based framework for real-time crash prediction on the basic freeway segments of urban expressways	X			
S17	A practical tool for public health surveillance: Semi-automated coding of short injury narratives from large administrative databases using Naïve Bayes algorithms				
S18	Risk analysis during tunnel construction using Bayesian Networks: Porto Metro case study		X		
S19	Near-miss narratives from the fire service: A Bayesian analysis		X		
S20	Bayesian random effect models incorporating real-time weather and traffic data to investigate mountainous freeway hazardous factors		X		

VI. CONCLUSIONES

A continuación, se presenta un resumen de las principales conclusiones a las que se llegaron por medio de la RSL:

- El primer aporte de este trabajo en el uso de redes bayesianas para el análisis de riesgos fue que se pudo determinar que la inferencia bayesiana es usada con mayor frecuencia para la toma de decisiones, al igual que el Manto o Mapeo de Markov, seguido del clasificador bayesiano.
- Se concluye que las redes bayesianas permiten de manera gráfica la evaluación probabilística en el análisis de riesgos, a su vez dichas redes

se pueden modelar con cualquier proceso probabilístico que se componga de relaciones causales, y al poderse diagramar y programar se puede llegar a la toma de decisiones de manera más intuitiva que es lo que permite analizar el riesgo en cada área.

- Podemos concluir que las redes bayesianas son una opción viable para aplicar en el análisis de riesgos en distintas áreas de estudio en las que existe un ambiente de incertidumbre y de información escasa o de calidad cuestionable.

REFERENCIAS

- [1] B. Kitchenham, "Procedures for Performing Systematic Reviews," *Jt. Tech. Rep.*, 2004.

- [2] J. Thomas, W. Reinartz, and V. Kumar, "Getting the Most out of All your Customers," vol. 82, no. 2004, pp. 2005–2007, 2004.
- [3] M. R. Lozano, "El papel de las redes bayesianas en la toma de decisiones," *Lab. Model. y Simulación*, no. 2, p. 11, 2011.
- [4] BASEL COMITÉ ON BANKING SUPERVISION, "Sound Practices for the Management and Supervision of Operational Risk," Publ. núm. 86., no. January, 2001.
- [5] D. Rodríguez and J. Dolado, "Redes Bayesianas en la Ingeniería del Software Introducción", Univ. Alcalá, Univ. del País Vasco, pp. 1–21, 2007.
- [6] Q. Zhang, Y. Xu, X. Xu, and Z. Lou, "Application of Bayesian network in water quality risk analysis and pollution reduction decision making from small data," 2011 International Conference on Remote Sensing, Environment and Transportation Engineering. pp. 84–88, 2011.
- [7] C. Cornalba, R. G. Bellazzi, and R. Bellazzi, "Building a Normative Decision Support System for Clinical and Operational Risk Management in Hemodialysis," vol. 12, no.5, pp. 678–686, 2008.
- [8] E. Bompard et al., "Information Impact on the Risk Analysis of the Malicious Attack against Power System," 2007.
- [9] L. Tao et al., "Operational risk assessment of distribution network with consideration of PV output uncertainties," no. Ciced, pp. 10–13, 2016.
- [10] G. F. Cooper, "The computational complexity of probabilistic inference using bayesian belief networks," *Artif. Intell.*, vol. 42, no. 2–3, pp. 393–405, 1990.
- [11] F. G. J. Absil, "A Risk-Based ject-Oriented Approach to Sensor Management," 2005.
- [12] D. Dong, W. Sun, Z. Lv, J. Li, and Z. Qian, "Evaluation system of water inrush in coalmine based on J2EE," 2010 2nd International Conference on Future Computer and Communication, vol. 3. pp. V3-759-V3-762, 2010.
- [13] Z. Chen, H. Huang, Y. Liu, L. He, and Z. Wang, "Maintainability Verification for Airplanes with Small Samples Based on Similarity Degree," pp. 11–13, 2011.
- [14] Z. Li, T. J. Oechtering, and M. Skoglund, "PRIVACY-PRESERVING ENERGY FLOW CONTROL IN SMART GRIDS , School of Electrical Engineering and the ACCESS Linnaeus Centre KTH Royal Institute of Technology , Stockholm , Sweden," 2016.
- [15] A. P. Douglas, A. M. Breipahl, F. N. Lee, S. Member, R. Adapa, and W. B. R. Norman, "Risk Due to Load Forecast Uncertainty in Short Term Power System Planning * School of Electrical and Computer Engineering , University of Oklahoma," vol. 13, no. 4, pp. 1493–1499, 1998.
- [16] A. H. Marshall, L. A. Hill, and F. Kee, "Continuous Dynamic Bayesian Networks for Predicting Survival of Ischaemic Heart Disease Patients," pp. 178–183, 2010.
- [17] R. S. Carvalho, R. N. Carvalho, M. Ladeira, E. M. Monteiro, and G. L. d. O. Mendes, "Using Political Party Affiliation Data to Measure Civil Servants' Risk of Corruption," 2014 Brazilian Conference on Intelligent Systems. pp. 166–171, 2014.
- [18] J. Biolchini, P. Gomes, A. Cruz, and G. Horta, "Systematic Review in Software Engineering," *Chem. Des. Autom. News (CDA News)*, vol. 7, no. 12, pp. 16–22, 1992.
- [19] S. E. Group and R. Unido, "Directrices para la realización sistemática de la literatura críticas en Ingeniería de Software Sección de Control de Documentos." 2007
- [20] R. Yu, M. Abdel-aty, and M. Ahmed, "Bayesian random effect models incorporating real-time weather and traffic data to investigate mountainous freeway hazardous factors," *Accid. Anal. Prev.*, vol. 50, pp. 371–376, 2013.
- [21] M. Hossain and Y. Muromachi, "A Bayesian network based framework for real-time crash prediction on the basic freeway segments of urban expressways," *Accid. Anal. Prev.*, vol. 45, pp. 373–381, 2012.

El reto de la Tecnología de la Información y Comunicación para la educación en el Ecuador

The challenge of Information and Communication Technology for education in Ecuador

Silvana Sayago Heredia

silvana_sayago@hotmail.com

Resumen:

En este artículo se presenta información relacionada con las Tecnologías de la Información y la Comunicación reconociendo su concepto, características, ventajas y riesgos, todo esto desde la perspectiva de la educación, el objetivo es documentar y analizar diferentes manuscritos y estudios sobre la incidencia de los medios informáticos, programas de capacitación en la educación de nuestro país. Se destaca las funciones de las TIC'S, el rol del docente en estos nuevos escenarios, los nuevos modelos constructivistas que da lugar a nuevas modalidades de estudio, por lo que es evidente que es un reto para este medio de enseñanza el cual potencializa la calidad educativa.

Para contar con información se recogen y agrupan un buen número de investigaciones, de ámbito nacional e internacional, está recogida de información trata de ser exhaustiva, pero esto no significa que no existan otras investigaciones y temáticas que no se han mencionado aquí, que no por ello son menos significativas sino que no es posible abarcarlas todas. Se concluye mediante este trabajo que las instituciones u organizaciones que administran la enseñanza como el Ministerio de Educación, universidades, centros de enseñanza, etc. Requieren nuevos sistemas de distribución de los materiales, nuevas estructuras de comunicación de usuario con la organización, en consecuencia, nuevas fórmulas de diseño y producción de los materiales y cursos, por lo que también se recomienda mayor capacitación docente en temas de tecnologías.

Palabras Clave: Tecnologías de la Información y Comunicación, educación, docentes.

Abstract: This article presents information related to Information and Communication Technologies recognizing its concept, characteristics, advantages and risks, all from the perspective of education, the objective is to document and analyze different manuscripts and studies on the incidence of computer media, training programs in education in our country. It highlights the

functions of ICTs, the role of the teacher in these new scenarios, the new constructivist models that leads to new modes of study, so it is clear that it is a challenge for this teaching medium which potentiates educational quality.

In order to have information, a good number of researches are gathered and grouped, nationally and internationally, the information is gathered to be exhaustive, but this does not mean that there are no other investigations and topics that have not been mentioned here, that they are less significant, but it is not possible to cover them all. It concludes by means of this work that the institutions or organizations that administer the education like the Ministry of Education, universities, educational centers, etc. They require new systems of distribution of materials, new structures of user communication with the organization, consequently, new formulas for design and production of materials and courses, so that further teacher training on technology issues is also recommended.

Keywords: Information Technology and Communication, education, teachers.

1. INTRODUCCIÓN:

En el mundo de hoy nace, crece y se educa una generación influida por la presencia de las tecnologías de la información y comunicación y de la relación entre ambas es decir lo que se denomina telemática, responsable en gran medida de un cambio de percepción de la realidad fundamentalmente entre los miembros de este grupo de personas en desarrollo, así la computadora e internet se presentan como las TIC'S más emblemáticas de la actualidad. La Tecnología de la Información y Comunicación (TIC), se han convertido en algo habitual puesto que los niños, niñas y jóvenes conviven con ellas desde que nacen por lo que es algo natural en sus vidas y pueden aportar elementos positivos en su desarrollo como: acceso a la información, incentivar la comunicación, colaboración y ampliar formas de diversión. En el presente artículo se describe la incidencia de las TIC'S en la educación para identificar algunos elementos que obligan a un cambio profundo en las prácticas docentes, considerando las nuevas maneras de aprender que hoy en día se están generando. Lo anterior, con la finalidad de reflexionar sobre los procesos educativos soportados en las TIC y

ampliar el conocimiento de los docentes y gestores de las instituciones educativas al respecto, logrando vislumbrar alternativas que permitan la formación de una ciudadanía que pueda “vivir, aprender y trabajar con éxito en una sociedad cada vez más compleja, rica en información y basada en el conocimiento” [1].

2. LAS NUEVAS TECNOLOGÍAS:

Primeramente se debe definir que son las TIC'S ya que este término abreviado se utiliza mucho para las Tecnologías de la Información y Comunicación las mismas que han permitido llevar la globalidad al mundo de la comunicación, facilitando la interconexión entre las personas e instituciones a nivel mundial y eliminando barreras espaciales. Se denominan Tecnologías de la Información y las Comunicación al conjunto de tecnologías que permiten la adquisición, producción, almacenamiento, tratamiento, comunicación, registro y presentación de informaciones, en forma de voz, imágenes y datos contenidos en señales de naturaleza acústica, óptica o electromagnética [2].

2.1 Características de las TIC'S:

- **Inmaterialidad:** Mediante la digitalización es posible almacenar grandes cantidades de información. A su vez los usuarios pueden acceder a información ubicada en dispositivos electrónicos lejanos, que se transmite utilizando las redes de comunicación, de una forma transparente e inmaterial. Esta característica, ha venido a definir lo que se ha denominado como "realidad virtual", esto es, realidad no real.

- **Instantaneidad:** Podemos transmitir la información instantáneamente a lugares muy alejados físicamente. Se han acuñado términos como ciberespacio, para definir el espacio virtual, no real, en el que se sitúa la información, al no asumir las características físicas del objeto utilizado para su almacenamiento, adquiriendo ese grado de inmediatez e inmaterialidad.

- **Aplicaciones Multimedia:** Las aplicaciones o programas multimedia han sido desarrollados como una interfaz amigable y sencilla de comunicación, para facilitar el acceso a las TIC'S de todos los usuarios. Una de las características más importantes de estos entornos es "La interactividad". Otra de las características más relevantes de las aplicaciones multimedia, y que mayor incidencia tienen sobre el sistema educativo, es la posibilidad de transmitir información a partir de diferentes medios: texto, imagen, sonido, animaciones, etc. [2].

2.2 Ventajas, Oportunidades y posibles Riesgos por el uso de las TIC'S:

Las TIC permiten múltiples oportunidades y beneficios entre los cuales tenemos:

- Favorecen las relaciones sociales.
- El aprendizaje cooperativo.
- Desarrollo de nuevas habilidades y destrezas.
- Nuevas formas de construcción del conocimiento.
- Desarrollo de capacidades de creatividad.
- Comunicación y razonamiento.
- Inclusión a las personas con discapacidad.

En el ámbito familiar al reconocer la ventaja tecnológica que tiene la juventud, sobre los adultos se abre un nuevo espacio de participación en la familia. Así las TIC, lejos de convertirse en un elemento de aislamiento, bien utilizadas podrían ser un canal de comunicación entre los miembros del hogar y no un motivo de disputa. Pero a la vez existe una postura negativa por posibles problemas que puede causar el uso inadecuado de las mismas, algunos aspectos que pueden tener efectos negativos son: adicción, aislamiento, contenidos inadecuados y/o violentos, pornografía. Las instituciones, padres de familia y docentes están dedicando empeño para que se dé un buen manejo de las TIC'S pues pueden perjudicar a los jóvenes en su desarrollo personal y social [3].

3. INFLUENCIA DE LAS NUEVAS TECNOLOGÍAS EN LA EDUCACIÓN:

La era "Internet" exige cambios en el ámbito educativo, y los profesionales de la educación involucrados tienen múltiples razones para aprovechar las nuevas posibilidades que proporcionan las TIC para impulsar este cambio hacia un nuevo paradigma educativo más personalizado y centrado en el estudiante. Además de la necesaria alfabetización digital de los alumnos y del aprovechamiento de las TIC'S para la mejora de la productividad en general, el alto índice de fracaso escolar y la creciente multiculturalidad de la sociedad con el consiguiente aumento de la diversidad del alumnado es decir migrantes, constituyen razones poderosas para aprovechar las posibilidades de innovación metodológica que ofrecen las TIC para lograr una escuela más eficaz inclusiva [4].

En la docencia tradicional el aula se entiende como el espacio físico donde las relaciones directas son entre profesor y alumno, así como entre alumnos. En la actualidad frente a este modelo tradicional emerge un nuevo modelo cuyas características más destacadas son las siguientes:

- El aula se concibe como un espacio virtual, donde se relacionan profesor y alumno, así como los propios alumnos entre sí, lo que permite eliminar las barreras geográficas existentes.
- Los discentes pueden adaptar los calendarios y horarios, fijando el ritmo de trabajo que ellos consideren más adecuado en función de sus necesidades. Esto permite de una relación asíncrona, donde no es necesario que profesor y alumno coincidan en el mismo tiempo.
- La presencia física de los profesores y alumnos no es precisa para el desarrollo de las clases.
- Los participantes no es necesario que coincidan en el mismo instante de tiempo, lo que ofrece una mayor flexibilidad a los

alumnos especialmente, en la fijación de sus propios horarios [3].

Las pizarras digitales, internet y ordenadores pueden mejorar la enseñanza, pues crean otra dinámica pedagógica y una mayor participación de los educandos en el proceso de enseñanza, mejorar la autoestima y el trabajo en equipo. El nuevo modelo de enseñanza que propician las nuevas herramientas tecnológicas implica transformaciones en la cultura misma por ende en el proceso educativo, en los objetivos educativos, infraestructura y cambios en los contenidos didácticos

- **Proceso Educativo:** En la actualidad el aprendizaje es constante, una persona debe continuar su aprendizaje a lo largo de toda su vida, las TIC'S favorecen la formación continua al ofrecer herramientas que permiten la aparición de entornos virtuales de aprendizaje.

- **Cambios en los Objetivos Educativos:** Los educadores y las instituciones educativas deben preparar a los alumnos para vivir en la Sociedad del Conocimiento o Sociedad de la Información, por lo tanto se deben plantear objetivos desde muy tempranas edades para que los alumnos potencialicen el uso de las TIC'S.

- **Infraestructura:** Para poder educar con TIC las instituciones educativas de todos los niveles deben contar con ordenadores y tener internet banda ancha a mas de una serie de equipos.

- **Cambio en los contenidos didácticos:** Frente a los tradicionales libros, vídeos y juegos, los nuevos contenidos educativos creados se hacen más interactivos más atractivos y más variados, los profesores tienen la oportunidad de generar contenidos educativos de acuerdo con los intereses a las particularidades de sus alumnos y de su contexto educativo [5].

3.1 Incidencia de las TIC en el ámbito educativo y la necesidad de una alfabetización digital.

Sabemos ya que las TIC pueden convertirse en instrumentos útiles para mejorar la calidad y eficiencia de los procesos educativos. Ello se debe a que ayudan a crear entornos de aprendizaje que promueven la creatividad e innovación de los estudiantes y docentes revolucionando la forma en que se obtiene, se maneja y se interpreta la información. Por tanto, las TIC pueden ir ayudando a avanzar “de la docencia, a aprendizaje autodirigido, y, del aprendizaje como un evento único, al proceso de aprendizaje durante toda la vida” [6].

Podemos afirmar entonces, que las nuevas tendencias en materia de educación contemplan a las TIC como un elemento de gran importancia que no se puede ignorar puesto que “la formación de las nuevas generaciones no puede quedar al margen de la sociedad digital y no se trata sólo de proporcionar acceso a las TIC sino de formar para una utilización adecuada” [7].

Es importante comprender el término alfabetización [7], la consideran como el “conjunto de habilidades socio-cognitivas mediante las cuales se puede seleccionar, procesar, analizar e informar del proceso de transformación de información a conocimiento”. La alfabetización digital no es un fin en sí misma, sino un medio para lograr que las personas adquiramos las capacidades necesarias para llegar a ser competentes en el uso de las TIC, y ello implica:

- Buscar, analizar y evaluar información.
- Solucionar problemas y tomar decisiones.
- Utilizar creativa y eficazmente herramientas de productividad.
- Comunicar, colaborar, publicar y producir.
- Contribuir a la sociedad, siendo ciudadanos informados y responsables.

Tabla 1. Ventajas y desventajas del modelo tradicional de educación y la utilización de las Tics

Modelo de educación	VENTAJAS	DESVENTAJAS
	Aula se concibe como un modelo virtual	Se requiere estar dotados de ordenadores
	Relaciones virtuales entre alumnos y profesores	Conexión a internet banda ancha
	Relaciones virtuales entre alumnos	Mal uso del internet por parte de los alumnos.
	Permite eliminar las barreras geográficas	Desconocimiento para la supervisión de trabajos académicos
	Se establecen calendarios dinámicos y horarios fijando el ritmo de trabajo	Falta de conocimiento de los docentes de mayor edad
	Educación a distancia	
LAS TICS	No precisa que alumno y profesor coincidan en el tiempo La pizarra cuaderno y bolígrafo se vincula al ordenador y proyector	
	Libros digitales	
	Libros digitales interactivos mas atractivos y variados	
	Se puede generar contenidos de acuerdo a los intereses de los alumnos	
	Docente y estudiantes realizan sus trabajos utilizando el internet por buscadores web	
	VENTAJAS	DESVENTAJAS
Tradicional	Relaciones directas entre el profesor y alumno	Limitados al espacio físico del aula

Sociabilización y relaciones directas entre alumnos	Clases con periodicidad establecida
Los alumnos reciben y abstraen las ideas transmitidas del profesor	Limitado a la pizarra al cuaderno y bolígrafo
	Libros tradicionales
	La clase a largo plazo carece de interés
	Precisa que ambos agentes coincidan en el tiempo

Fuente: Las nuevas tecnologías en la educación [3].

Gráfico 1. Modelo tradicional y con uso de las TIC'S



Fuente: Samsung y el IPN crean aula virtual para desarrolladores [8].

4. LAS TIC'S PARA LA EDUCACIÓN ECUATORIANA:

En Ecuador se detectó un número importante de proyectos, pero los de mayor cobertura fueron los siguientes que aportaron a la educación en el país:

✓ **Maestr@s.com:** es un proyecto del Ministerio de Educación y Cultura de Ecuador; su objetivo principal es: “mejorar y lograr aprendizajes en niños y jóvenes a través de la incorporación de TIC'S en el trabajo docente de sus maestros y en la utilización de las tecnologías para la elaboración de proyectos de aula o institucionales que den como resultado una mayor calidad de la educación” [9]. Mediante el programa se otorga un estímulo económico en dólares y líneas de crédito blando a los docentes beneficiarios para la adquisición de equipo informático y capacitación en el manejo de las computadoras y en sus aplicaciones educativas.

✓ **Educación Ecuador:** portal educativo del Ministerio de Educación y Cultura.

✓ **Programa Nacional de Tele-educación:** está impulsado por el Ministerio de Educación y Cultura junto con el Consejo Nacional de Telecomunicaciones y varias universidades, sus objetivos son mejorar la calidad de la educación mediante la incorporación de

proyectos de tele-educación; desarrollar una infraestructura de telecomunicaciones; desarrollar contenidos nacionales y culturales para Internet; y poner especial atención en la utilización de las tecnologías informáticas como herramientas para promover el desarrollo de habilidades superiores de pensamiento [10],

✓ **Proyecto Navegar:** pertenece a la Unión Nacional de Educadores para dar capacitación a maestros en el área de informática básica.

Por otra parte, las universidades ecuatorianas también realizan un gran esfuerzo para apoyar la incorporación y uso de las TIC en la educación. Entre los proyectos más importantes destacan los siguientes:

✓ **Escuela Superior Politécnica del Litoral (ESPOL):** entre sus proyectos sobresalen dos:

- **Innovación de la Educación Pública Primaria en la Península de Santa Elena (IEPSE),** que busca desarrollar la creatividad y las habilidades cognitivas de los estudiantes de las escuelas primarias públicas de la península a través de la introducción de las TIC en las aulas de zonas rurales y urbanas marginales.

- **Red Latinoamericana de Educación (RELATED),** cofinanciado por la Fundación Ford, que pretende mejorar la educación en los niveles de educación secundaria y superior en Latinoamérica mediante la utilización de las TIC.

✓ **Universidad Técnica Particular de Loja (UTPL):** es pionera en el área de la educación a distancia y, a juicio de algunos de los entrevistados, con sus aulas virtuales y desarrollo de programas de cómputo, es el ejemplo más preciso de aplicación de las TIC dentro del ámbito universitario [11],

4.1 Los profesores, las TIC y los procesos de Enseñanza:

Nuevas formas de enseñanza en la educación ecuatoriana es precisa pues como manifiesta [12], “el profesor deja de ser fuente de todo conocimiento y el profesor pasa a actuar de guía de alumnos para facilitarles el uso de recursos y herramientas que necesitan para explorar y elaborar nuevo conocimiento, pasa a actuar como gestor” [13]. Pues el docente al igual que el sistema educativo evoluciona en la aplicación de modelos de aprendizaje generando material acorde al contexto. El reto del profesorado consistirá en ser competente en el manejo adecuado y pedagógico de las TIC y transformar las metodologías tradicionales en estrategias innovadoras que promuevan la construcción de aprendizajes.

5. CONCLUSIONES:

❖ Es preciso promover en el profesorado habilidades y competencias para aprovechar al máximo las posibilidades de la

tecnología en el contexto educativo, adaptándola e integrándola al proceso de enseñanza-aprendizaje, es decir guiar a los alumnos en el uso de las bases de información y conocimiento así como proporcionar acceso a los alumnos para usar sus propios recursos.

❖ En nuestro país se convierte en un reto alfabetizar a la sociedad para el manejo eficiente de medios tecnológicos pues se sigue invirtiendo en programas de capacitación para el uso de las TIC'S, en especial para docentes con programas de formación continua, también para la comunidad en general con aulas virtuales en diferentes instituciones públicas, siendo una fuerte inversión del estado para formar una nueva sociedad acorde a los cambios.

❖ Cuando se reflexiona acerca de las ventajas de las TIC'S, la primera es aquella referente a un modelo nuevo de aprendizaje que rompe barreras de tiempo, espacio y da lugar a que una persona con discapacidad sea incluida, pero también tiene riesgos ya que la web es un medio de comunicación masivo y de fácil acceso por lo que propicia negocios ilícitos, creando grupos vulnerables, así las instituciones educativas tienen que la obligación de guiar para que este medio sea una herramienta para el desarrollo de la república.

REFERENCIAS BIBLIOGRÁFICAS:

- [1] Cumbre Mundial sobre la sociedad de la información- Ginebra (2003). Construir sociedades de la información que atiendan a las necesidades humanas, Agencia Latinoamericana de Información Recuperado del sitio web: <https://www.itu.int/wsis/docs2/tunis/off/6rev1-es.html>.
- [2] Rosario, Jimmy, (2005). "La Tecnología de la Información y la Comunicación (TIC). Su uso como Herramienta para el Fortalecimiento y el Desarrollo de la Educación Virtual". Disponible en: <http://www.cibersociedad.net/archivo/articulo.php?art=218> [visitado en octubre de 2014]
- [3] Moya, A, (2009). "Las nuevas tecnologías en la educación" [En línea], Granada, disponible en: http://www.csi-csif.es/andalucia/modules/mod_ense/revista/pdf/Número_24/ANTONIA_M_MOYA_1.pdf [visitado en octubre de 2014]
- [4] Marqués, P. (2012). "Impacto de las TIC'S en la educación: funciones y Limitaciones", Revista de Investigación Barcelona, disponible en: <http://www.3ciencias.com/wp-content/uploads/2013/01/impacto-de-las-tic.pdf> [visitado en octubre de 2014]
- [5] Laboarda, R. (2005). ""Las nuevas tecnologías en la educación", Fundación AUNA, Madrid. [En línea], disponible en: http://www.universidadcentral.cl/prontus_ucentral/site/artic/20091204/asocfile/20091204170522/texto_8.pdf [visitado en septiembre de 2014]
- [6] Unesco (1997). Alfabetización y Tecnología. Hamburgo: Unesco-Quinta Conferencia Internacional de Educación de las Personas Adultas.
- [7] Gross, B. & Contreras, D. (2006). La alfabetización digital y el desarrollo de competencias ciudadanas. Revista Iberoamericana de Educación, (42), pp. 103-125.
- [8] Samsung y el IPN crean aula virtual para desarrolladores, <https://expansion.mx/tecnologia/2016/08/16/samsung-y-el-ipn-crean-aula-virtual-para-desarrolladores>. [visitado noviembre 2018]
- [9] Salinas, J. (1997). "Nuevos ambientes de aprendizaje para una sociedad de la información", Edutec, disponible en: <http://www.ses.unam.mx/curso2008/pdf/Salinas.pdf>
- [10] Valdivieso, Tania. (2010). "Uso de tic en la práctica docente de los maestros de educación básica y bachillerato de la ciudad de Loja", Revista electrónica de Tecnología Educativa, disponible en: http://edutec.rediris.es/Revelec2/Revelec33/pdf/Edutec-e_n33_Salome.pdf [visitado en octubre de 2014]
- [11] Vidal, Mª.P. (2006). Investigación de las TIC en la educación, Revista Latinoamericana de Tecnología Educativa, 5 (2), 539-552. http://www.unex.es/didactica/RELATEC/sumario_5_2.htm [Visitado en octubre de 2014]
- [12] Ramírez, J. (2005). Las tecnologías de la información y de la comunicación en la educación en cuatro países latinoamericanos, Redalyc, disponible en: <http://www.redalyc.org/pdf/140/14002805.pdf>. pp. 9,10.
- [13] La pizarra y el pizarrón, <https://loquequierenospeques.blogspot.com/2016/06/al-pizarron.html>. [visitado noviembre 2018]

SISTEMA SCADA PARA MONITOREO DE CONDICIONES FÍSICO Y QUÍMICAS DE UN ACUARIO DE PECES TROPICALES

SCADA SYSTEM FOR MONITORING THE PHYSICAL AND CHEMICAL CONDITIONS OF A TROPICAL FISH AQUARIUM

Marlon Cedeño, Carolina Burbano, Jenniffer Bautista, Mariana Pachay, Manuel Nevárez

Colaboradores del Laboratorio de Investigación, Tecnologías e Innovación – LITI, Escuela de Sistemas y Computación, PUCE Esmeraldas, Ecuador {[marlon.cedeno](mailto:marlon.cedeno@pucese.edu.ec), [carolina.burbano](mailto:carolina.burbano@pucese.edu.ec), [jenniffer.bautista](mailto:jenniffer.bautista@pucese.edu.ec), [mariana.pachay](mailto:mariana.pachay@pucese.edu.ec)}@pucese.edu.ec
 Coordinador del Laboratorio de Investigación, Tecnologías e Innovación – LITI, Escuela de Sistemas y Computación, PUCE Esmeraldas, Ecuador, manuel.nevarez@pucese.edu.ec

Resumen. Al momento de implementar un acuario, se deben considerar varios factores, ya que se pretende imitar a la naturaleza creando un ecosistema no natural de similares características al hábitat de los peces, entre los parámetros que intervienen en el desarrollo de las especies se presentan la temperatura, pH, oxígeno disuelto en el agua (ODA) y nivel de iluminación. El proyecto consiste en implementar un sistema de monitoreo y control de condiciones ambientales fisicoquímicas, que permitan simular un ecosistema favorable para la vida de estas especies. Las tendencias en automatización apuntan hacia la creación de sistemas de control con aplicaciones SCADA, la comunicación de datos entre el prototipo y el sistema se realizó con un módulo Ethernet Shield y Protocolo de Internet (IP) para la transmisión de datos a la base de datos del servidor, la aplicación web se controlará desde un sitio llamado "Squarium" manteniendo una constante interacción entre el usuario y el sistema de monitoreo, mostrando las variaciones de los sensores y permitiendo rapidez en la toma de decisiones, esto produjo un ambiente óptimo para los peces y la disminución del mantenimiento del acuario.

Palabras Clave: control, automatización, SCADA, peces tropicales, acuario, Arduino Ethernet.

Abstract. When implementing an aquarium, several factors must be considered, since it is intended to imitate nature by creating a non-natural ecosystem of similar characteristics to fish habitat, among the parameters that intervene in the development of the species presented the temperature, pH, dissolved oxygen in the water (ODA) and level of illumination. The project consists of implementing a system of monitoring and control of physicochemical

environmental conditions that simulate a favorable ecosystem for the life of these species. The trends in automation point towards the creation of control systems with SCADA applications, the data communication between the prototype and the system was made with an Ethernet Shield module and Internet Protocol (IP) for the data transmission to the server database, the web application will be controlled from a site called "Squarium" maintaining a constant interaction between the user and the monitoring system, showing the variations of the sensors and allowing rapid decision-making, this produced an environment optimal for fish and decreased maintenance of the aquarium.

Keywords: control, automation, SCADA, tropical fish, aquarium, Arduino Ethernet.

1 Introducción

La ciencia que estudia el entorno de los acuarios de la denomina Acuarofilia, se cree que los primeros en desarrollar esta afición fueron los egipcios, quienes criaron peces en estanques por razones místicas y ornamentales [1]. La Acuarofilia, se define como la afición a la cría de peces en un acuario, bajo condiciones controladas como son: la temperatura, iluminación, y pH del agua. Algunos complementos tecnológicos importantes en un acuario son la iluminación de alta intensidad y espectro continuo [2].

En un acuario tropical, el agua debe mantener una temperatura entre rangos de 23 y 28 °C aproximadamente, se mantiene dentro de este rango con un termo calentador, diseñado específicamente para el acuario requerido [3].

Las nuevas tendencias en automatización apuntan hacia la implementación de sistemas de control, como son aplicaciones del SCADA, un software que tiene como finalidad la comunicación con el usuario a través de una interfaz hombre máquina, es importante que un sistema de control de procesos interactúe con el

usuario, esta interacción va a permitir monitorear el proceso y a la vez asignar los valores respectivos [4]. De esta manera, el clásico supervisor soportado por un sistema de automatización es concebido como un sistema de control que integre las tareas de detección y diagnóstico de fallas, como una actividad previa que permite incorporar de manera natural el control tolerante a fallas [5].

VB (Visual Basic) en conjunto con la plataforma libre Arduino, permiten el desarrollo de sistemas para el control, adquisición y supervisión de datos. El aplicar un control electrónico a un acuario, facilita su mantenimiento y preservación de especies acuáticas, mediante el empleo de una tarjeta de adquisición de datos y un software SCADA, el usuario tiene el control del acuario, de esta manera puede informarse sobre las variables como la temperatura, oxigenación, iluminación, entre otras, con el objetivo de permitirle al usuario tomar la decisión correcta. El sistema de automatización basado en un entorno web, hace uso de la tecnología IP, es decir permite la comunicación a través de la red, obteniendo como resultado el control remoto del acuario [6].

También existen sistemas de control para la preservación de peces en acuarios desarrolladas en otros tipos de plataformas Open Source como es Raspberry, Open Aquarium es un ejemplo de ello, el usuario puede visualizar los datos de los sensores a través de una pantalla táctil, los mismos que se utilizan para ser visualizados en una aplicación web y un Streaming del acuario. La implementación de estos proyectos da como resultado un sistema electrónico de bajo costo y un control total sobre su acuario de un modo sencillo y eficaz [7].

Uno de las tecnologías que ofrece Arduino para recrear comunicación de datos es el módulo de comunicación Xbee, compuesto por un transmisor y un receptor, que usan poca energía para funcionar, diseñado para aplicaciones que requieren comunicaciones seguras con una mínima tasa de datos [8], se presenta como alternativa a uso de cables. Este módulo se encarga de realizar comunicación inalámbrica a través del protocolo Zigbee, tienen un rango de transmisión entre los 100 y 120 metros (Xbee regular) a espacio abierto, dependiendo de su serie y modelo [9].

Los datos de los sensores se transmitirán a un módulo de Arduino llamado Ethernet Shield, para mantener la información actualizada en la página web "Squarium", este módulo permite conectar fácilmente Arduino a Internet por medio de la librería <Ethernet.h> para enviar y recibir datos desde

cualquier parte del mundo con una conexión a Internet. Se pueden implementar sistemas de control de forma remota desde un sitio web [10].

2 Materiales y Métodos

El Proyecto consiste en el desarrollo de un sistema de control automático que integre a la plataforma Open Source Arduino, para monitorizar y controlar la información recibida por diferentes sensores ambientales, para la preservación de un ecosistema artificial acuático para peces tropicales. Se toma como referencia el sistema "Aquasys" implementado anteriormente [1], donde se definen las variables que están relacionadas a las condiciones fisicoquímicas de un acuario como son el pH, Oxígeno disuelto en el agua (ODA), nivel de iluminación y temperatura, para mantener el ecosistema interior estable. La finalidad es recrear ambientes subacuáticos, utilizando materiales electrónicos, sensores ambientales y software libre.

Se inició con la identificación de señales provenientes de los diferentes sensores, que interactúan con en el acuario a la placa de arduino. El sensor de temperatura, iluminación y pH emite una señal a la placa. Luego de interconectar los diferentes sensores, se realizaron las pruebas de funcionamiento del prototipo, para la medición de pH, se utilizaron patrones NIST de calibración con valores $\pm 0,02$ unidades de pH a 25 °C, para la medición temperatura se utilizó una sonda DS18B20 que soporta valores entre 24 °C – 28°C, el esquema de conexión de los dispositivos se muestra en la figura 1.

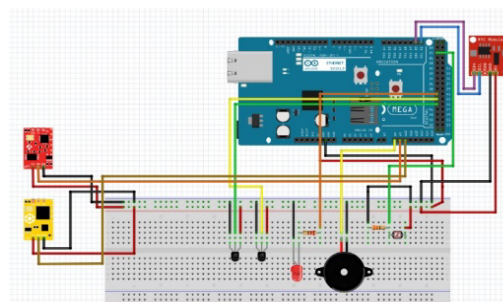


Figura 1. Circuito del sistema "Squarium", se muestran las conexiones del sistema de monitoreo de señales del acuario. Se inicializan la variable del sensor de temperatura en los pines (42-44), el pH (A9), ODA (A10), iluminación (24), reloj SDA(20), SCL(21) del Arduino MEGA para que los datos sean mostrados en la página web.

La función del módulo de control es recoger la información proporcionada por los sensores y actuar en consecuencia, si la temperatura es mayor de 28°C se activará una alarma en el sistema indicando que debe regularse la temperatura, como solución se apagarán las luces del acuario y encenderá la ventilación del agua de retorno para que normalice la temperatura ambiente del acuario. Los datos de muestran en una pantalla LCD para que sean interpretadas por el usuario. La transmisión de los valores de los sensores son almacenados en la base de datos "Squarium" creada

en MySQL, alojado en servidor local. Se hace uso del Ethernet Shield, puesto que por medio del protocolo de internet (IP) se logra establecer la comunicación a través de la red obteniendo como resultado final el control automático del acuario.

3 Resultados y Discusión

En la página web “Squarium”, desarrollada en código PHP, que muestra una tabla con la información de los sensores, la temperatura en grados centígrados, la misma que debe de estar entre los 24°C – 28°C, el pH en nivel de purificación 4 -10, la iluminación en 256 unidades lux y el nivel de ODA en 5-7% de Saturación, la información obtenida se muestra en la figura 2.



Figura 2. Página de inicio del sistema Squarium, se muestran la página de inicio del sistema Squarium, se visualizan los valores de temperatura, ODA, pH, e Iluminación. En la página principal se encuentran los enlaces para los diferentes reportes de los sensores y el manual de usuario del sistema.

El sistema controla que los valores se mantengan en los intervalos establecidos, configuración para la conexión del Arduino con el servidor se muestra a continuación:

```
byte mac[] = {0xDE, 0xAD, 0xBE, 0xEF, 0xEE, 0xFF}; //
Direccion MAC
byte ip[] = { 192,168,11,20 }; // Direccion IP del Arduino
byte server[] = { 192,168,1,54};
EthernetClient client;
////////////////////////////////////
Serial.println("Connecting...");
if (client.connect(server, 80)>0) { // Conexion con el
servidor
  client.print("GET
/Sarduino/models/servidor.php?t1="); // Enviamos los datos
por GET
  client.print(sensors1.getTempCByIndex(0));
  client.print("&t2=");
  client.print(sensors2.getTempCByIndex(0));
  client.print("&ilumi=");
  client.print(valorLDR);
  client.print("&dia=");
  client.print(now.day());client.print("/");
  client.print("&mes=");
  client.print(now.month());client.print("/");
```

```
client.print("&anio=");
client.print(now.year());
client.print("&ph=");
client.print(phValue);
client.print("&oda=");
client.print(Valor_O2,1);
client.println(" HTTP/1.0");
client.println("User-Agent: Arduino 1.0");
client.println();
Serial.println("Conectado");
} else {
  Serial.println("Fallo en la conexion");
}
if (!client.connected()) {
  Serial.println("Disconnected!");
}
client.stop();
client.flush();
delay(3000); // Espero un minuto antes de tomar otra muestra
```

A continuación se muestra parte del código php utilizado para insertar los datos provenientes de sensores, en la base de datos *Squarium*, los valores obtenidos se muestran en la figura 3.

```
<?php
require("base.php");
$valor = mysqli_real_escape_string($con, $_GET['fecha']);
$valor1 = mysqli_real_escape_string($con, $_GET['t1']);
$valor2 = mysqli_real_escape_string($con, $_GET['t2']);
$valor3 = mysqli_real_escape_string($con, $_GET['ilumi']);
// fecha reloj arduino
$valor4 = mysqli_real_escape_string($con, $_GET['dia']);
$valor5 = mysqli_real_escape_string($con, $_GET['mes']);
$valor6 = mysqli_real_escape_string($con, $_GET['anio']);
$valor7 = mysqli_real_escape_string($con, $_GET['ph']);
$valor8 = mysqli_real_escape_string($con, $_GET['oda']);
$query = "INSERT INTO
sensors(fecha,t1,t2,ilumi,dia,mes,anio,ph,oda) VALUES
('".$valor."','".$valor1."','".$valor2."','".$valor3."','".$valor4."','".$valor5."',
'".$valor6."','".$valor7."','".$valor8."')";
// Ejecutamos la instrucción
mysqli_query($con, $query);
mysqli_close($con);
?>
```

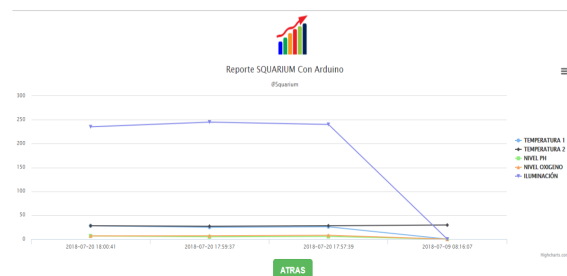


Figura 3. Reporte sensores general de los sensores, el gráfico presenta en tiempo real de la temperatura del agua de retorno y el agua dentro del acuario, pH, ODA e Iluminación se, un dato general de todos los sensores.

Conclusiones.

Se comprobó el correcto funcionamiento del sistema, los valores de los sensores de temperatura se encuentran en los rangos predeterminados de 22°C a 28°C y acorde a las condiciones físicas se debe controlar, esta no debe superar 28°C, en trabajos anteriores este control de temperatura no superaba los 32°C [11]. Los valores de pH en agua sucia son de 7.8 a 8 unidades pH y los niveles lux de iluminación de deben superar 256 lux.

Cada cinco segundos se actualizan los valores de los sensores en la base de datos y se visualizan en el sistema web SCADA Squarium, logrando de este modo mantener al usuario constantemente informado del estado de los factores fisicoquímicos que conforman el ecosistema natural del acuario para la toma de decisiones de manera automática, permite mantener un ecosistema confiable para la preservación de especies, posibilitando brindar un ambiente de vida controlado y disminuir la posibilidad de contaminación de este ecosistema.

El hardware y software abierto permitieron crear una solución innovadora, técnica, confiable y económica para realizar el control continuo de la calidad de agua en el acuario. La aplicación del sistema de control automático de las condiciones fisicoquímicas de un acuario de peces tropicales "Squarium", reduce el trabajo de mantenimiento del acuario.

Referencias Bibliográficas.

- [1] C. Guerrero, "Sistema computarizado para el control automático de las condiciones físico químicas de un acuario de peces tropicales," Pontificia Universidad Católica Del Ecuador Sede En Esmeraldas, 2014.
- [2] F. Lango, M. Castaneda, J. Zamora, G. Hernandez, M. Ramirez, and E. Solis, "La acuariofilia de especies ornamentales marinas: un mercado," *Latin American Journal of Aquatic Research*, 2012. [Online]. Available: https://scielo.conicyt.cl/scielo.php?pid=S0718-560X2012000100002&script=sci_arttext.
- [3] D. Chicaiza and H. Flores, "Parámetros biológicos de *Pseudocurimata boulengeri* (Characiformes: Curimatidae) en el embalse Chongón, Ecuador," *Rev. Biol. Trop.*, 2016.
- [4] J. Herrera, M. Barrios, and S. Pérez, "Diseño e implementación de un sistema scada

- inalámbrico mediante la tecnología zigbee y arduino Desing and implementation of a wireless scada system by means of zigbee and arduino technology," vol. 12, no. 2, pp. 45–50, 2014.
- [5] M. Cerrada, J. Cardillo, and A. Prada, "Diagnóstico de fallas basado en modelos: Una solución factible para el desarrollo de aplicaciones SCADA en tiempo real Model-based fault diagnosis: A feasible approach to develop SCADA applications in real time," *Rev. Cienc. e Ing.*, vol. 32, no. 323, pp. 163–172, 2011.
 - [6] N. D. J. VÁZQUEZ, "Scada para control de acuarios mediante arduino y vb," *Escuela Universitaria Politécnica*, 2015.
 - [7] V. Lapuente, "control y monitorización de un acuario en tiempo real mediante tecnología open source," *Universidad de Zaragoza Escuela de Ingeniería y Arquitectura*, 2014.
 - [8] C. Durán and H. García, "Desarrollo de un Sistema Inalámbrico para la Supervisión y Control de un Aerogenerador Development of a Wireless System for Monitoring and Control of a Wind Turbine Durán & García / Desarrollo de un Sistema Inalámbrico para la Supervisión y Control de un Aerogenerador," *Tecno Lógicas, Medellín*, pp. 123–7799, Oct-2013.
 - [9] A. R. Chung Pinzás and E. Ruiz Lizama, "Diseño e implementación de un software scada para el módulo de temperatura del," *Revista de la Facultad de Ingeniería Industrial*, pp. 42–46, Jun-2011.
 - [10] Y. Mulge, "Remote Temperature Monitoring Using LM35 sensor and Intimate Android user via C2DM Service," *IJCSMC*, vol. 2, no. 6, pp. 32–36, 2013.
 - [11] A. J. Correa Vivanco, "Sistema de control electrónico para acuarios utilizando tecnologías gsm y voip," *Universidad Técnica de Ambato Facultad de Ingeniería en Sistemas Electrónica e Industrial Carrera de Ingeniería en Electrónica y Comunicaciones*, 2017.

Método de seguridad en la autenticación de contraseña utilizando AES

Security method in password authentication using AES

Andrade B. Alicia, Herrera F. Boris

(Universidad Central del Ecuador, Ciudadela Universitaria Av. América, Quito, aandrade@uce.edu.ec, bherrera@uce.edu.ec)

Resumen— El algoritmo AES (Advanced Encryption Standard) y las funciones hash criptográficas se emplean en el área de la criptografía con propiedades especializadas, siendo uniformes y con efecto avalancha, además dotadas de parámetros ajustables que permiten fallar ataques por fuerza bruta o de diccionario sobre las bases de datos que almacenan el resumen de las contraseñas de los usuarios, incluso cuando estos atacantes emplean hardware especializado.

Nuestra propuesta plantea una función propia incorporando el algoritmo AES y como resultado un generador pseudoaleatorio de contraseñas, en varias pruebas basadas en espacio de memoria y tiempo. En base a los resultados obtenidos en Lenguaje GO, se los compara en el mismo lenguaje y con los mismos parámetros utilizando las conocidas funciones hash de contraseñas como: Argon2 y Scrypt (Key derivation functions), obteniendo resultados positivos de mejora dado que se explota la ventaja que tiene AES de ser acelerado por hardware en la mayoría de los procesadores modernos, finalmente se realizará la evaluación con proyección de graficas de los resultados de la simulación y del desempeño de las funciones hash criptográficas.

Palabras clave. — *PBKDF, Cryptography, AES, Hash*

Abstract.— The AES algorithm (Advanced Encryption Standard) and the cryptographic hash functions are used in the area of cryptography with specialized properties, being uniform and with avalanche effect, also equipped with adjustable parameters that allow failures by brute force or dictionary on the databases that store the summary of users' passwords, even when these attackers use specialized hardware.

Our proposal proposes a proper function incorporating the AES algorithm and as a result

a pseudorandom generator of passwords, in several tests based on memory space and time. Based on the results obtained in GO Language, they are compared in the same language and with the same parameters using the known hash functions of passwords like: Argon2 and Scrypt (Key derivation functions), obtaining positive results of improvement given that the advantage that AES has of being accelerated by hardware in the majority of modern processors, finally the evaluation will be carried out with projection of graphs of the results of the simulation and the performance of the cryptographic hash functions.

Keywords— *PBKDF, Cryptography, AES, Hash*

I. INTRODUCCIÓN

Los algoritmos utilizados están pensados principalmente como funciones password-hashing, con la función principal de autenticación del usuario proporcionando algunas contramedidas contra atacantes a GPUs rápidas o hardware personalizado.

La gestión de las contraseñas de los usuarios no es un tema menor y, generalmente, se procesan mediante una función hash criptográfica junto con una cadena aleatoria conocida como sal.

El uso de unidades de procesamiento gráfico (GPUs) e, incluso, hardware específico ha tenido un gran impacto, permitiendo ataques por fuerza bruta

a las bases de datos que contienen los hashes de las contraseñas. Por ello, ha sido necesario utilizar funciones de hash de contraseña especiales, con parámetros de espacio y tiempo ajustables, para hacer más costosos los ataques por fuerza bruta [8].

Las funciones de derivación de claves basadas en contraseña (password based key derivation functions, PBKDF) es un campo de investigación muy activo, existiendo una gran cantidad de publicaciones relacionadas con estas funciones de hash de contraseña [2,8,12] que mejoran el conocido estándar PBKDF2

[9].

El algoritmo AES ([5]) es un elemento básico en la criptografía actual. Aunque por naturaleza, es más lento que algunos cifrados de flujo actuales,

```
fH := sha3.New256()
fH.Write(pass)
fH.Write(0)
fH.Write(salt)
seed := fH.Sum(nil)

blk :=
aes.NewCipher(seed[0:16])
fC :=
cipher.NewCTR(blk, seed[16:32])

out := make([]byte, plen)
fC.XORKeyStream(out, out)

M := make([]byte, pmem*plen)
fC.XORKeyStream(M, M)
```

pero tiene la ventaja de ser acelerado por hardware en muchos procesadores modernos, lo que hace que sea extremadamente rápido, con velocidades superiores a 1 GB/s en sistemas actuales. Es una excelente primitiva básica sobre la que construir muchas herramientas criptográficas, incorporando en su diseño la velocidad y seguridad que aporta este estándar. En este trabajo se describe una PBKDF (o función de hash de contraseñas) [1] basada en el empleo de AES en modo contador (CTR) como un generador de números pseudoaleatorios (o pseudo-random number generator, PRNG).

Analizamos su rendimiento y seguridad, y lo comparamos en términos de rendimiento en una cantidad equivalente de uso de memoria con las funciones PBKDF Scrypt [10] y Argon2[4].

La función de hash de contraseñas propuesta puede ser especialmente útil en aplicaciones donde se requiera autenticación segura y sea importante tener un rendimiento configurable, como en dispositivos móviles, computación embebida, Internet de las cosas, redes de sensores remotas, entre otros.

En lo que respecta a las dimensiones de seguridad, son 4 las consideradas:

Autenticidad: Propiedad o característica consistente en que una entidad es quien dice ser o bien que garantiza la fuente de la que proceden los datos.

Confidencialidad: Propiedad o característica consistente en que la información ni se pone a disposición, ni se revela a individuos, entidades o procesos no autorizados.

Disponibilidad: Propiedad o característica de los

activos consistente en que las entidades o procesos autorizados tienen acceso a los mismos cuando lo requieren.

Integridad: Propiedad o característica consistente en que el activo de información no ha sido alterado de manera no autorizada.

II. DESCRIPCIÓN

En este proyecto se toma como base pruebas anteriores, tomando la función y en comparación con Scrypt y Argon2, se emplearon las implementaciones proporcionadas en los paquetes oficiales de golang.org/x/crypto.

La llamada de función y los parámetros son los mismos que en la función original.

A. Parámetros:

`pass[]`, array de bytes de longitud arbitraria que contiene la contraseña (password) del usuario.

`salt[]`, array de bytes de longitud arbitraria que contiene la sal aleatoria. Se recomienda un mínimo de 32 bytes.

`plen`, entero positivo que indica la longitud (length) en bytes del hash de salida. Se recomienda un mínimo de 32 bytes.

`pmem`, entero positivo que indica el número de entradas en la tabla principal. Este parámetro modula el coste espacial (memoria). Se requiere un valor mínimo de 1.

`ptime`, entero positivo que indica el número de pasadas a través de la tabla principal. Este parámetro modula el coste temporal (time). El valor mínimo posible es 1.

B. Elementos

`M[]`, array de `pmem` x `plen` bytes de longitud. Constituye el bloque de memoria principal del algoritmo.

`out[]`, array de `plen` bytes que contiene el hash de salida. `fH()`, una función hash criptográfica segura con un tamaño de resumen de 256 bits utilizado para la semilla. En nuestro caso, empleamos el estándar SHA3-256.

`fC()`, un cifrador en bloque seguro, en modo CTR. En nuestro caso, empleamos el estándar AES-128 en modo CTR. Es la primitiva criptográfica básica en el diseño de la función propuesta y se utiliza principalmente como un PRNG.

C. Inicialización de la función

En el lenguaje de pseudocódigo empleado (basado en lenguaje Go [6]), se asume que se puede llamar a `len()` en arrays y obtener su longitud, evitando tener que pasar longitudes como argumentos separados.

```
func
Key(pass, salt []byte, plen, pmem, ptime
int)
```

Para evitar colisiones en las que la concatenación directa de la contraseña y la sal produzcan como resultado los mismos datos de entrada, la semilla inicial se obtiene como el hash de la contraseña y la sal separadas por un solo byte con valor 0. De ese hash de 256 bits, los primeros 16 bytes de la semilla se usan como la clave de 128 bits para AES y los segundos 16 bytes de la semilla conforman el valor de inicialización de 128 bits para el modo CTR. A continuación, los búferes `out[]` y `M[]` se cifran utilizando la instancia de AES inicializada. Se supone que estos almacenamientos intermedios son puestos a cero por `make()` antes del cifrado. A continuación, se muestra el pseudocódigo para la etapa de inicialización:

D. Salida

El algoritmo de salida es sencillo ya que, para cada entrada en `M[]`, se obtiene un índice pseudoaleatorio de los primeros 8 bytes (64 bits) de `out[]` que, a su vez, se emplea para elegir una entrada de `M[]` a combinar con el estado actual de `out[]`. Esta operación consiste en una resta de bytes (Z_{256}). Finalmente, el resultado de esta operación se cifra a `out[]`. El ciclo externo es el número de veces (o pasadas por `M[]`) que se realiza todo este proceso, siendo la salida de la función el último estado de `out[]`. Dado que `out[]` es la salida directa de un cifrado AES, toda la función hash de contraseña debe ser tan segura como AES. A continuación, se muestra el pseudocódigo para la etapa de salida:

```

for t := 0; t < ptime; t++ {
  for m := 0; m < pmem; m++ {
    i := (int(out[0:8])%pmem) * plen
    for o := 0; o < len(out); o++ {
      M[i+o] -= out[o]
    }
    fC.XORKeyStream(out, M[i:i+plen])
  }
}

```

Tabla1. Inicialización para la etapa de salida. Original (AESCTR-o)

```

for t := 0; t < ptime; t++ {
  for m := 0; m < pmem; m++ {
    i := (int(out[0:8])%pmem) * plen
    for o := 0; o < len(out); o++ {
      M[i+o] -= out[o]
    }
    i = (i*i) % pmem
    for o := 0; o < len(out); o++ {
      out[o] -= (M[i+o] ^ out[o])
    }
  }
}
fC.XORKeyStream(out, out)

```

Tabla2. Inicialización para la etapa de salida. Intermedia (AESCTR-i)
Las diferencias entre las optimizaciones finales e intermedias se resaltan en azul en la Tabla 4.

```

for t := 0; t < ptime; t++ {
  for m := 0; m < pmem; m++ {
    i := (int(out[0:8])%pmem) * plen/8
    for o := 0; o < len(out); o++ {
      out64[o] -= (M64[i+o] ^ out64[o])
    }
  }
}
fC.XORKeyStream(out, out)

```

Tabla3. Inicialización para la etapa de salida. Final (AESCTR-f)

III. ANALISIS

A. Rendimiento

A continuación, analizamos las características de rendimiento y seguridad de nuestra propuesta y comparamos las variantes con `Scrypt` y `Argon2`. Las pruebas experimentales se han implementado en el lenguaje de programación `Go` (versión 1.10) y se han llevado a cabo en una computadora con Intel Core i7 a 2,90GHz (que admite aceleración de hardware AES) y 16 GB de RAM; tanto las contraseñas y sales empleadas han sido de 32 bytes (256 bits) de longitud, al igual que la salida, que también ha sido de 32 bytes; cada prueba se midió diez veces, tomando el mínimo como el valor final.

En la figura 1, (realizada y proyectada en EXCEL, con los datos generados en GO) se muestra el coste computacional de la función

propuesta al aumentar el parámetro pmem exponencialmente de 2^8 a 2^{23} entradas de 32 bytes (lo que equivale a un uso de memoria desde 8KiB a 256MiB) y con una sola pasada (ptime = 1). Se observa que se obtienen valores muy razonables de alrededor de 0,5 segundos o menos con pmem hasta 2^{21} , en las tres variantes del pseudocódigo. También es destacable que existe un margen significativo disponible, en caso de desear una mayor complejidad computacional.

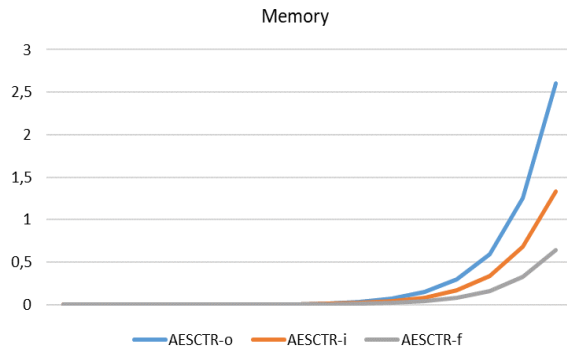


Figura 1. Rendimiento en función de la complejidad espacial con las 3 versiones de AES en modo CTR (en segundos)

En este caso, la diferencia en el rendimiento de la optimización final (AESCTR-f) es más pronunciada que en el caso del parámetro de memoria.

El rendimiento del parámetro ptime se muestra en la figura 2, (realizada y proyectada en EXCEL, con los datos generados en GO). En este caso, ptime aumenta exponencialmente de 1 a 2^{15} pasadas y pmem se fija en 256 entradas de 32 bytes (que equivale a 8KiB de memoria). Se obtienen tiempos de ejecución razonables de alrededor de 0,5 segundos o menos con hasta 2^{14} pasadas

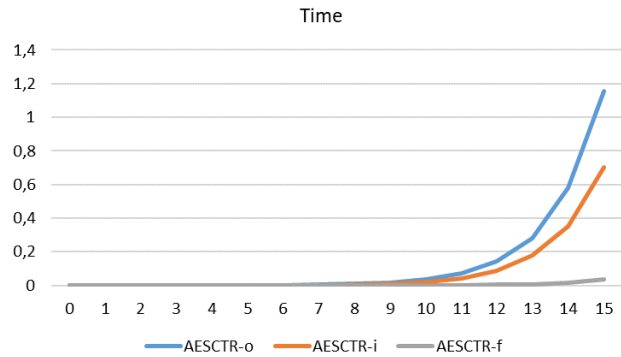


Figura 2. Rendimiento en función de la complejidad temporal (en segundos)

La figura 3, (realizada y proyectada en EXCEL, con los datos generados en GO) muestra el tiempo de ejecución asociado a la modulación simultánea en un doble bucle de los parámetros pmem y ptime. El número de entradas en M[] (bucle externo) va de 2^8 a 2^{15} y el número de pasadas (bucle interno) va de 1 a 2^7 .

La serie esperada de curvas exponenciales crecientes se observa claramente, mostrando un buen rango de rendimiento en la modulación de ambos parámetros de complejidad. El rendimiento también es muy razonable dentro de estos parámetros, con todos los tiempos alrededor de 0,5 segundos o menos. La cantidad máxima de memoria procesada es de 128MiB con $pmem=2^{15}$ y $ptime=2^7$

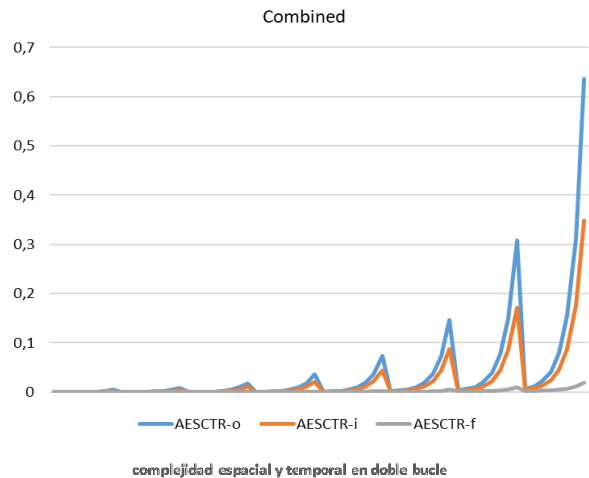


Figura 3. Rendimiento en función de ambas complejidades combinadas en segundos

1) Comparación con Argon2 y Scrypt

Una de las funciones de hash de contraseñas utilizada más en los servicios actuales de autenticación mediante contraseña es la PBKDF Argon2 [9], que ha sido diseñada para defenderse

de los ataques mediante GPU y permite utilizar una cantidad configurable de memoria. Se puede observar en la figura 4 que la función propuesta en este trabajo (basada en AES-CTR) es significativamente más rápida que Argon2 al procesar cantidades equivalentes de memoria. Para obtener los resultados mostrados en la gráfica, se ha hecho uso de la librería estándar de Go que implementa AES (que aprovecha la aceleración de hardware) y la de Argon (que no está acelerada por hardware y utiliza la versión de ronda reducida de Salsa20/8 en su núcleo). El resultado obtenido pone de manifiesto que AES es una buena opción para diseñar funciones de hash de contraseñas en la mayoría de las máquinas modernas; por el contrario, en plataformas donde AES no este acelerado por hardware, el rendimiento será considerablemente menos competitivo.

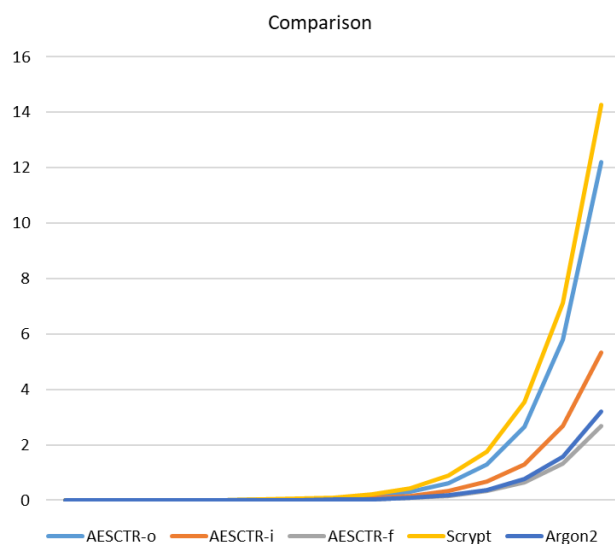


Figura 4. Comparativa de las 3 versiones de AESCTR propuestas con Scrypt y Argon2

B. Seguridad

El algoritmo propuesto está diseñado, fundamentalmente, como función de hash de contraseñas con un objetivo principal: la autenticación de usuarios mediante contraseña; al tiempo, proporciona algunas contramedidas contra ataques que hacen uso de unidades de procesamiento gráfico rápidas o de hardware especializado. Las principales trabas que interpone contra este tipo de ataques son el uso de memoria (evitando, así paralelismo) y el tiempo requerido para obtener el hash, explotando el hecho de que AES se implementa en hardware en la mayoría de los procesadores modernos [4,6,11,13,14]; estas trabas no las ofrecen las funciones hash

criptográficas comunes, que se ejecutan en estas plataformas especializadas mucho más rápido (varios órdenes de magnitud), permitiendo ataques por fuerza bruta satisfactorios.

Otra aplicación que se le puede dar es convertir una contraseña de longitud variable en una clave de longitud variable, con excelentes propiedades de avalancha o también como un esquema de prueba de trabajo en los protocolos con cadenas de bloques (*blockchain*) usados en las criptomonedas. En la función propuesta, la primitiva criptográfica principal es AES-128 [5] en modo CTR, analizado para 3 versiones, lo que permite una alta seguridad y rendimiento en la mayoría de sistemas modernos.

La clave y el valor de inicialización para AES-CTR se obtienen haciendo uso del hash SHA3-256 [5] de la contraseña y la sal proporcionadas.

Tanto AES como SHA3 son primitivas que se consideran seguras y estándares actuales ampliamente empleados; el diseño propuesto debería estar acorde con estas primitivas en términos de seguridad, proporcionando un nivel de seguridad mínimo de 128 bits para prevenir ataques por fuerza bruta. En el caso de que SHA3-256 o AES-128 se rompieran, podrían utilizarse otras primitivas de seguridad equivalentes en su lugar.

IV. CONCLUSIONES

La función de hash de contraseñas con tres versiones presentada basada en Advanced Encryption Standard en modo contador que aprovecha el soporte de hardware acelerado en la mayoría de sistemas modernos para prevenir ataques por fuerza bruta para la recuperación de contraseñas que emplean hardware especializado o unidades de procesamiento gráfico de propósito general.

Se ha analizado el rendimiento de dicha función propuesta en varios escenarios de memoria y tiempo y se ha realizado un estudio comparativo con un estándar de facto dentro de las funciones de hash de contraseñas como Argon2 y Scrypt, obteniéndose resultados prometedores y demostrando que AES es un candidato adecuado para este tipo de primitivas criptográficas.

La seguridad general de la función diseñada es equivalente a la de AES, ya que se utiliza como una primitiva central en el diseño y la salida final es el resultado directo del cifrado AES.

A futuras investigaciones, podría ser muy interesante realizar una implementación en hardware especializado para probar de forma adecuada las características de rendimiento en estas plataformas. Además, se podría considerar en el contexto de la función propuesta otras medidas de seguridad adicionales como el uso de un archivo aleatorio grande en el servidor (*server-side ROM*), poder actualizar la seguridad sin intervención del usuario final (*client-independent update*) o delegar parte del trabajo al usuario final (*server relief*), así como la evaluación comparativa del rendimiento con otras propuestas recientes.

REFERENCIAS

- [1] Alvarez, R., Andrade, A., Santos, I., Zamora, A.: AES-CTR as a Password-Hashing Function. *Advances in Intelligent Systems and Computing*, vol. 649, 610-617 (2018)
- [2] Alvarez, R., Zamora, A.: Using Spritz as a Password-Based Key Derivation Function, en *Advances in Intelligent Systems and Computing*, vol. 527, pp. 518–525, 2017.
- [3] Bertoni, G., Daemen, J., Peeters, M., Van Assche, G.: Cryptographic sponge functions. <http://sponge.noekeon.org/>, 2011.
- [4] Biryukov, A., Dinu, D., Khovratovich, D.: Argon2: the memory-hard function for password hashing and other applications”, *Password Hashing Competition winner*, 2016. <https://github.com/P-H-C/phc-winner-argon2/blob/master/argon2specs.pdf>
- [5] Daemen, J., Rijmen, V.: AES proposal: Rijndael. In *First Advanced Encryption Standard (AES) Conference (1999)*
- [6] Forler, C., Lucks, S., Wenzel, J.: The Catena Password-Scrambling Framework. Version 3.2, *Bauhaus-Universität Weimar*. (2015) <https://www.uni-weimar.de/fileadmin/user/fak/medien/professuren/Mediensicherheit/Research/Publications/catena-v3.2.pdf>
- [7] The Go Programming Language. <http://www.golang.org>
- [8] Hellman, M. E.: A cryptanalytic time-memory trade-off. *Information Theory, IEEE Transactions on*, vol. 26–4, 401–406. (1980)
- [9] Kaliski, B.: PKCS #5: Password-Based Cryptography Specification Version 2.0. Internet Engineering Task Force, Network Working Group, Request for Comments (RFC) 2898 (2000) <https://tools.ietf.org/html/rfc2898#section-5.2>
- [10] Percival, C.: Stronger key derivation via sequential memory-hard functions. *BSDCan – The BSD Conference*. (2009) http://www.bsdcn.org/2009/schedule/attachments/87_scrypt.pdf
- [11] Pornin, T.: The MAKWA Password Hashing Function. Version 1.1. *Password Hashing Competition finalist*. (2015) <https://www.bolet.org/makwa/makwa-spec-20150422.pdf>
- [12] Provos, N., Mazieres, D.: A Future-Adaptable Password Scheme. *USE- NIX Annual Technical Conference, FREENIX track*, 81–91. (1999)
- [13] Simplicio, M. A., Almeida, L. C., Andrade, E. R., dos Santos, P.C.F., Barreto, P. S. L. M.: Lyra2: Password Hashing Scheme with improved security against time-memory trade-offs. *IACR Cryptology ePrint Archive*, 2015:136. (2015)
- [14] Solar Designer: yescrypt – password hashing scalable beyond bcrypt and scrypt. Presented at PHDays’14. Openwall (2014) <http://www.openwall.com/presentations/PHDays2014-Yescrypt/PHDays2014-Yescrypt.pdf>

Análisis comparativo de metodologías ágiles de desarrollo de software: una revisión bibliográfica

Comparative analysis of agile software development methodologies: a bibliographic review

Xavier Quiñónez-Ku, Juan Casierra Cavada, Luis Herrera-Izquierdo, Junior Mera Quiroz

Carrera de Sistemas y Computación, Pontificia Universidad Católica del Ecuador Sede Esmeraldas,
xavier.quinonez@pucese.edu.ec,

Carrera de Sistemas y Computación, Pontificia Universidad Católica del Ecuador Sede Esmeraldas,
juan.casierrac@pucese.edu.ec,

Professional Services, ThoughtWorks EC, lherrera@thoughtworks.com,
 LatAm Autos, jmera@latamautos.com.

Resumen— El concepto de metodologías ágiles ha generado un gran interés a las organizaciones para lograr objetivos de desarrollo de software rápidos y funcionales, gracias a sus características de concentrarse más en la entrega de un software funcional en lugar de presentar una carga masiva de documentación; respuesta rápida a los cambios en los requisitos en lugar de seguir un plan prescrito; colaboración con clientes en lugar de negociación de contratos y dando más preferencia a individuos e interacciones sobre procesos y herramientas. En esta investigación se realizó una comparativa entre las metodologías ágiles a través de una revisión bibliográfica, para ayudar en la selección de la metodología de desarrollo de software apropiada en un escenario particular.

Abstract-- The concept of agile methodologies has generated great interest for organizations to achieve fast and functional software development objectives, thanks to their characteristics of concentrating more on the delivery of a functional software instead of presenting a massive documentation load; quick response to changes in requirements instead of following a prescribed plan; collaboration with clients instead of negotiating contracts and giving more preference to individuals and interactions about processes and tools. In this research, a comparison was made between the agile methodologies through a bibliographic review, to assist in the selection of the appropriate software development methodology in a particular scenario.

I. INTRODUCCIÓN

Es común encontrar personas cuyo criterio acerca del desarrollo de software está ligado únicamente a los programas de computadora, lo cual es bastante erróneo, ya que el desarrollo de

software también está relacionado con todos aquellos documentos que recaban información acerca de los requerimientos, el diseño y demás procesos vinculados [1].

El desarrollo de software no es una tarea sencilla, por mucho tiempo esta labor se ha llevado adelante sin una metodología definida. La disciplina que se encarga de determinar, analizar, describir y establecer las diferentes metodologías y lineamientos para mejorar u optimizar el desarrollo de sistemas informáticos se la denomina Ingeniería de Software [2].

Algunos autores definen una metodología como una colección de procedimientos, técnicas, herramientas y documentos auxiliares que ayudan a los desarrolladores de software en sus esfuerzos por implementar nuevos sistemas de información. Una metodología está formada por fases, cada una de las cuales se puede dividir en sub-fases, que guiarán a los desarrolladores de sistemas a elegir las técnicas más apropiadas en cada momento del proyecto y también a planificarlo, gestionarlo, controlarlo y evaluarlo. Sin embargo, una metodología es algo más que una colección, puesto que se basa en una filosofía, distinguiéndose de los métodos o de las simples recetas, que marcan simplemente unos pasos a seguir. Así, las metodologías difieren ya sea por la cantidad de fases, las técnicas de cada fase, el contenido de la fase o en su base filosófica, todo esto se aplica, dependiendo del contexto de desarrollo, tamaño del proyecto o del equipo de trabajo, cultura organizacional, entre otros aspectos [3].

Según [4], la complejidad de los sistemas actuales, los cambios repentinos en el contexto de estos sistemas y las modificaciones en los requerimientos del cliente una vez que se ha comenzado el desarrollo de un proyecto de software, generan un ambiente donde la planificación, el desarrollo, la administración y el control del mismo resultan difíciles de estimar o evaluar. Este tipo de escenarios requiere de metodologías de desarrollo de software que permitan generar resultados rápidamente.

II. HISTORIA DEL ARTE

Metodologías de desarrollo

Según [5], en la década de los noventa surgieron metodologías de desarrollo de software ligeras, más conocidas como metodologías ágiles, que buscaban reducir la probabilidad de fracaso por subestimación de costos, tiempos y funcionalidades en los proyectos de desarrollo de software. Estas metodologías nacieron como reacción a las metodologías existentes con el propósito de disminuir la burocracia que implica la aplicación de las metodologías tradicionales en los proyectos de pequeña y mediana escala. Las metodologías tradicionales buscan imponer disciplina al proceso de desarrollo de software y de esa forma volverlo predecible y eficiente. Para conseguirlo se soportan en un proceso detallado con énfasis en planeación propio de otras ingenierías. El principal problema de este enfoque es que hay muchas actividades que hacer para seguir la metodología y esto retrasa la etapa de desarrollo.

Diferentes investigadores comparan enfoques tradicionales y ágiles en sus diferentes perspectivas, en la Tabla I se describen las diferencias entre las metodologías ágiles de desarrollo y las metodologías tradicionales según [6]–[10].

TABLA I
DIFERENCIAS ENTRE METODOLOGÍAS TRADICIONALES Y
ÁGILES

	Metodologías tradicionales	Metodologías ágiles
Hipótesis fundamental	Los sistemas son totalmente especificables, predecibles y se desarrollan a través de una planificación detallada y extendida.	El software adaptativo de alta calidad es desarrollado por pequeños equipos que utilizan el principio de mejora continua del diseño y las pruebas basadas en una rápida respuesta y cambio.
Estilo de gestión	Comando y control	Liderazgo y colaboración
Conocimiento administrativo	Explícito	Tácito
Comunicación	Formal	Informal
Modelo de desarrollo	Modelo de ciclo de vida (cascada, espiral o modelos modificados)	Modelo evolutivo de entrega
Estructura organizacional	Mecánico (burocrático, alta)	Orgánico (flexible y participativo,

	formalización), dirigido a grandes organizaciones.	fomenta la cooperación social), dirigido a pequeñas y medianas organizaciones.
Control de calidad	Planificación difícil y control estricto. Pruebas difíciles y tardías	Control permanente de requisitos, diseño y soluciones. Pruebas permanentes.
Requisitos de usuario	Detallado y definido antes de la codificación / implementación.	Entrada interactiva
Costo de reinicio	Alto	Bajo
Dirección de desarrollo	Fijo	Fácilmente cambiabile
Pruebas	Después de completar la codificación	Cada iteración
Participación del cliente	Bajo	Alto
Habilidades adicionales requeridas de los desarrolladores	Nada en particular	Habilidades interpersonales y conocimientos básicos del negocio.
Escala apropiada del proyecto	Gran escala	Baja y mediana escala
Desarrolladores	Planificado, con capacidades adecuadas, acceso a conocimientos externos	Ágil, con conocimientos avanzados y cooperativos.
Clientes	Con acceso al conocimiento, cooperativo, representativo y empoderado	Dedicado, informado, cooperativo, representativo y capacitado.
Requerimientos	Muy estable, conocido de antemano	Emergente, con cambios rápidos.
Arquitectura	Diseño para requerimientos actuales y previsibles.	Diseño para requerimientos actuales.
Remodelación	Costoso	No es caro
Tamaño	Grandes equipos y proyectos	Pequeños equipos y proyectos
Objetivos principales	Alta seguridad	Valor rápido

Nota: La tabla muestra las principales diferencias entre las metodologías ágiles y las metodologías tradicionales

Metodologías ágiles de desarrollo

El enfoque de las metodologías ágiles está teniendo una amplia efectividad en proyectos donde los requisitos son muy cambiantes,

ya que en este tipo de proyectos la comunicación con el cliente debe ser fundamental y precisamente ese es uno de los principios básicos de las metodologías ágiles.

En marzo de 2001, 17 críticos de los modelos de producción basados en procesos resumieron en cuatro postulados lo que hoy en día se conoce como el Manifiesto Ágil [11]. Los cuatro postulados mencionados en aquella reunión mencionaban los siguientes criterios: Individuos e interacciones sobre procesos y herramientas, Software que funciona sobre documentación exhaustiva, Colaboración de clientes sobre la negociación del contrato y Respuestas a cambios sobre seguir un plan. Estos cuatro postulados son considerados como la definición canónica del desarrollo ágil.

A continuación, en la Figura 1 se puede evidenciar las principales razones por las que las organizaciones a nivel mundial están adoptando las metodologías ágiles de desarrollo de software:



Fig. 1. Razones para adoptar metodologías ágiles [12].

En la actualidad existen un sin número de metodologías ágiles, unas más populares que otras, cada una aportando al desarrollo ágil distintos métodos que ayudan a mejorar de una manera eficaz la calidad del software. Entre las metodologías ágiles más populares se encuentran las siguientes:

A. Programación Extrema (XP):

La Programación Extrema o XP (Extreme Programming) es un enfoque de desarrollo de sistemas que acepta lo que se conoce como buenas prácticas en esta área y las lleva al extremo [13], [14]. En esta metodología cabe resaltar la importancia del cliente, las pruebas, la

refactorización, la simplicidad, la propiedad colectiva del código que se ven reflejadas en las cuatro prácticas esenciales de XP [15], [16]:

- *Entregas limitadas o pequeñas*: Consiste en realizar entregas parciales de módulos del sistema.
- *Semana de trabajo de 40 horas*: Los equipos de desarrollo de XP trabajan de manera intensa durante una semana típica de 40 horas.
- *Cliente en el sitio*: Esta práctica insiste en que el cliente debe hacer parte fundamental y activa del grupo de trabajo y debe estar presente durante todo el proceso de desarrollo.
- *Programación en Pareja*: Con esto se busca aumentar la calidad del código, ahorrar tiempo, estimula la creatividad y la reducción de código fuente.

La programación extrema engloba un conjunto de reglas y prácticas que ocurren en el contexto de cuatro actividades estructurales: planeación, diseño, codificación y pruebas. La Figura 2 muestra cada una de las actividades de XP, y resalta las tareas claves de cada una [13], [15], [17], [18]:



Fig. 2. El proceso de la programación extrema [15].

Planeación: Esta actividad comienza escuchando a los clientes, para entender el contexto del negocio y definir las características principales y funcionalidad que se requiere, estas características se transforman en requerimientos del negocio que se especifican mediante Historias de Usuario; las cuales recogen la interacción hablada entre desarrolladores y usuarios. Una vez hechas las Historias de Usuario, el equipo de desarrollo las divide en tareas, estima el esfuerzo, recursos requeridos para su implementación, se genera el plan de entregas, las iteraciones, la rotación de parejas y las reuniones diarias.

Diseño: Es la etapa en donde son evaluadas las historias de usuario por el equipo del proyecto para dividir las en tareas, cada tarea representa una característica distinta del sistema y se puede

diseñar una prueba de unidad que verifique cada tarea, estas tareas se representan por medio de las tarjetas CRC (Clase-Responsabilidad-Colaborador). Las tarjetas CRC identifican y organizan las clases bajo el paradigma orientado a objetos (lo que incluye asignación de responsabilidades), cada tarjeta contiene el nombre de la clase (que representa una o más historias de usuario), una descripción de las responsabilidades o métodos asociados con la clase, así como la lista de las clases con que se relaciona o que colaboran con ella. Las tarjetas CRC son el único trabajo de diseño que se genera como parte del proceso de XP.

Codificación: Se lleva a cabo la programación en pareja, la unidad de pruebas y la integración del código. Durante esta etapa se espera la disponibilidad del cliente para que éste pueda resolver cualquier duda que se presente durante una jornada de trabajo.

Prueba: Cada tarea que se identificó con las historias de usuario, representa una característica distinta del sistema y se realiza una prueba de unidad por cada una de ellas, existen pruebas unitarias las cuales son diseñadas para probar cada uno de los métodos y clases, dichas pruebas son realizadas por los programadores.

B. Scrum:

Su nombre no corresponde a una sigla, sino a un concepto deportivo, propio del rugby, relacionado con la formación requerida para la recuperación rápida del juego ante una infracción menor [5], [19]. Scrum es un modelo de desarrollo ágil caracterizado por [11], [20], [21]:

- Adoptar una estrategia de desarrollo incremental, en lugar de la planificación y ejecución completa del producto.
- Basar la calidad del resultado más en el conocimiento tácito de las personas en equipos autoorganizados, que en la calidad de los procesos empleados.
- Solapamiento de las diferentes fases del desarrollo, en lugar de realizarlas una tras otra.

Según [15], [22], los principios Scrum son congruentes con el manifiesto ágil y se utilizan para guiar actividades de desarrollo dentro de un proceso de análisis que incorpora las siguientes actividades estructurales: requerimientos, análisis, diseño, evolución y entrega. El flujo general del proceso Scrum se ilustra en la Figura 3:

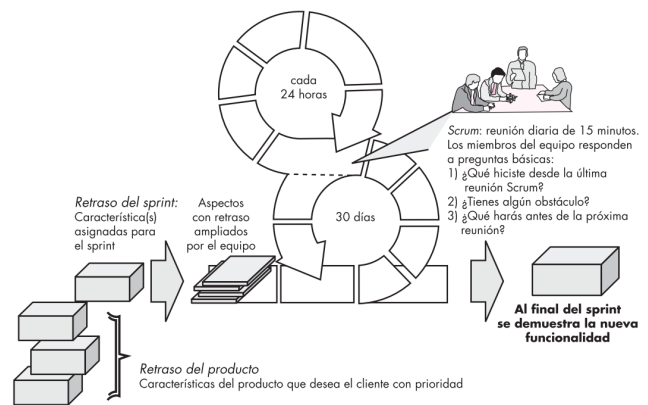


Fig. 3. Flujo del proceso Scrum [15].

Retraso: lista de prioridades de los requerimientos o características del proyecto que dan al cliente un valor del negocio.

Sprints: consiste en unidades de trabajo que se necesitan para alcanzar un requerimiento definido en el retraso que debe ajustarse en una caja de tiempo predefinida (lo común son 30 días).

Reuniones Scrum: son reuniones breves (de 15 minutos, por lo general) que el equipo Scrum efectúa a diario.

Demostraciones preliminares: entregar el incremento de software al cliente de modo que la funcionalidad que se haya implementado pueda demostrarse al cliente y éste pueda evaluarla.

Las características del desarrollo basado en Scrum son [10], [23]:

Colaboración: El desarrollo basado en Scrum promueve la colaboración ya que está impulsado por equipos multifuncionales donde cada persona con sus habilidades y experiencia contribuye a la mejor solución de diseño. Un equipo multifuncional incluye una combinación de programadores, arquitectos de software, analistas de software y expertos en control de calidad.

Reuniones diarias: La metodología Scrum está marcada por reuniones diarias de corta duración en las que el equipo de desarrollo de productos se comunica y evalúa el estado de progreso del desarrollo de software, lo que aumenta la productividad de los miembros del equipo.

Product Backlog: El product backlog captura los requisitos para que un producto de software se entregue con éxito. Mantiene una lista ordenada de características, correcciones de errores, requisitos no funcionales.

Sprint Backlog: El sprint backlog registra la lista de tareas que realizará el equipo de desarrollo durante el próximo sprint. Esta lista se elabora recogiendo las tareas desde la parte superior de la cartera de productos hasta que se realice el trabajo suficiente para el próximo sprint, teniendo en cuenta la capacidad de trabajo y los resultados pasados del equipo de desarrollo.

Roles: El desarrollo basado en Scrum se rige por 3 funciones principales [19], [24]:

- **Product Owner:** responsable de definir, priorizar y comunicar los requisitos del producto y guía el proceso de desarrollo del producto.

- Equipo de desarrollo: responsable de ejecutar las tareas asignadas por el propietario del producto dentro del plazo del sprint. Por lo general, un equipo multifuncional de 3 a 9 individuos implementa las tareas de desarrollo del producto previstas por el propietario del producto.
- Scrum Master: responsable de hacer cumplir las reglas y los principios del desarrollo basado en Scrum. El Scrum Master elimina los impedimentos para el desarrollo y ayuda a mejorar el proceso, el equipo de desarrollo y el producto de software que se está desarrollando.

C. Desarrollo adaptativo de software (ASD):

El desarrollo adaptativo de software (ASD) es una técnica para elaborar software y sistemas complejos. Los fundamentos filosóficos del ASD se centran en la colaboración humana y en la organización propia del equipo [15], [25], [26]. La Figura 4 ilustra el ciclo de vida del ASD:

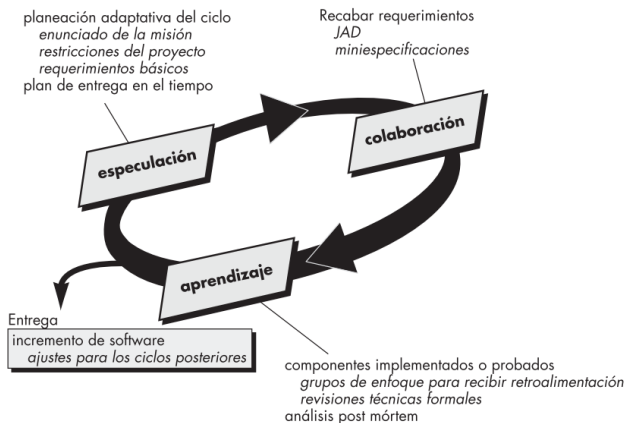


Fig. 4. Desarrollo adaptativo de software [15].

El “ciclo de vida” del ASD incorpora tres fases: especulación, colaboración y aprendizaje [15], [25], [27].

En la *especulación*, se inicia el proyecto y se lleva a cabo la planeación adaptativa del ciclo. La especulación emplea la información de inicio del proyecto para definir el conjunto de ciclos de entrega (incrementos de software) que se requerirán para el proyecto.

Las personas motivadas usan la *colaboración* de manera que multiplica su talento y producción creativa más allá de sus números absolutos. Este enfoque es un tema recurrente en todos los métodos ágiles. Sin embargo, la colaboración no es fácil. Incluye la comunicación y el trabajo en equipo, pero también resalta el individualismo porque la creatividad individual desempeña un papel

importante en el pensamiento colaborativo.

Conforme los miembros de un equipo ASD comienzan a desarrollar los componentes que forman parte de un ciclo adaptativo, el énfasis se traslada al *aprendizaje* de todo lo que hay en el avance hacia la terminación del ciclo.

D. Feature Drive Development (FDD):

El desarrollo impulsado por las características es una metodología ágil para el desarrollo de sistemas, basado en la calidad del software, que incluye un monitoreo constante del proyecto. Al igual que otros proyectos ágiles, adopta una filosofía que [15], [27], [28]:

- pone el énfasis en la colaboración entre los integrantes del equipo;
- administra la complejidad de los problemas y del proyecto con el uso de la descomposición basada en las características, seguida de la integración de incrementos de software, y
- comunica los detalles técnicos en forma verbal, gráfica y con medios basados en texto.

El desarrollo impulsado por las características se divide en 5 fases y es de tipo incremental, constando cada incremento (iteración) en 2 fases: diseño y construcción de una característica. En la Figura 5 se pueden distinguir estas fases:

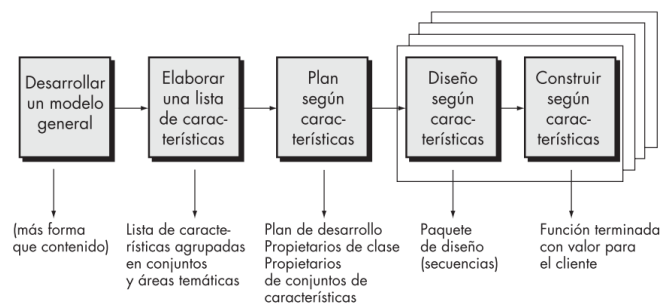


Fig. 5. Desarrollo impulsado por las características (Feature Drive Development (FDD)) [15], [25].

Se explica a continuación cada etapa [15], [25], [27]:

- *Desarrollar un modelo global*: Al inicio del desarrollo se construye un modelo teniendo en cuenta la visión, el contexto y los requisitos que debe tener el sistema a construir. Este modelo se divide en áreas que se analizan detalladamente. Se construye un diagrama de clases por cada área.
- *Construir lista de características*: Se elabora una lista que resuma las funcionalidades que debe tener el sistema, cuya lista es evaluada por el cliente. Cada funcionalidad de la lista se divide en funcionalidades más pequeñas para un mejor entendimiento del sistema.
- *Planificar*: Se procede a ordenar los conjuntos de funcionalidades conforme a su prioridad y dependencia, y se asigna a los programadores jefes.
- *Diseñar*: Se selecciona un conjunto de funcionalidades de la lista. Se procede a diseñar y construir la funcionalidad

mediante un proceso iterativo, decidiendo que funcionalidad se van a realizar en cada iteración. Este proceso iterativo incluye inspección de diseño, codificación, pruebas unitarias, integración e inspección de código.

- **Construir:** Se procede a la construcción total del proyecto.

E. Dynamic Systems Development Method (DSDM):

El método de desarrollo de sistemas dinámicos (DSDM) es un enfoque de desarrollo ágil de software que “proporciona una estructura para construir y dar mantenimiento a sistemas que cumplan restricciones apretadas de tiempo mediante la realización de prototipos incrementales en un ambiente controlado de proyectos” [11], [15]. El DSDM es un proceso iterativo de software en el que cada iteración sigue la regla de 80 por ciento. Es decir, se requiere sólo suficiente trabajo para cada incremento con objeto de facilitar el paso al siguiente. Los detalles restantes se determinan más tarde, cuando se conocen los requerimientos del negocio y se han pedido y efectuado cambios [9], [15], [29].

El ciclo de vida DSDM, define tres ciclos iterativos distintos, precedidos de dos actividades adicionales al ciclo de vida [30]:

Estudio de factibilidad: establece los requerimientos y restricciones básicas del negocio, asociados con la aplicación que se va a construir, para luego evaluar si la aplicación es un candidato viable para aplicarle el proceso.

Estudio del negocio: establece los requerimientos e información funcional que permitirán a la aplicación dar valor al negocio; asimismo, define la arquitectura básica de la aplicación e identifica los requerimientos para darle mantenimiento.

Iteración del modelo funcional: produce un conjunto de prototipos incrementales que demuestran al cliente la funcionalidad. El objetivo de este ciclo iterativo es recabar requerimientos adicionales por medio de la obtención de retroalimentación de los usuarios cuando practican con el prototipo.

Diseño e iteración de la construcción: revisa los prototipos construidos durante la iteración del modelo funcional a fin de garantizar que en cada iteración se ha hecho ingeniería en forma que permita dar valor operativo del negocio a los usuarios finales; la iteración del modelo funcional y el diseño e iteración de la construcción ocurren de manera concurrente.

Implementación: coloca el incremento más reciente

del software en el ambiente de operación.

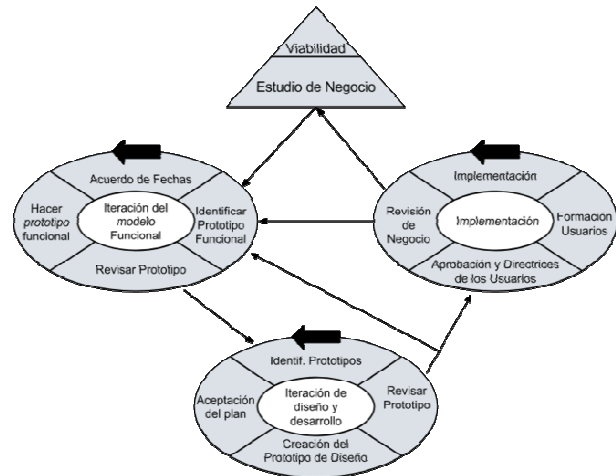


Fig. 6. Ciclo de vida DSDM [30].

III. ANÁLISIS COMPARATIVO ENTRE METODOLOGÍAS DE DESARROLLO ÁGILES

Aunque existen varias metodologías de desarrollo ágiles disponibles, cada una se aplica a un conjunto específico de proyectos. Un proyecto de desarrollo de software tiene varios factores asociados, como el tamaño del proyecto, la complejidad, el tiempo asignado, el presupuesto, etc. La selección de la metodología adecuada para el desarrollo de software depende de tales factores. Por lo tanto, un análisis comparativo de metodologías ágiles ayudará a decidir cuál se puede utilizar en una situación dada [31], [32].

A. Documentación:

Uno de los principios del desarrollo ágil es reducir la cantidad de tiempo y esfuerzo dedicado a la documentación. Pero la documentación que es importante no puede ser eliminada por completo [33], [34]:

- En metodologías como Scrum, XP y ASD la documentación es de menor importancia.
- En proyectos con FDD se requiere más documentación.
- DSDM requiere un nivel moderado de documentación que aún es menor que FDD.

B. Interacción con el cliente:

Las metodologías ágiles otorgan una importancia primordial a la comunicación frecuente con los usuarios finales. Aun así, el grado de implicación es diferente en cada metodología [35], [36]:

- XP y Scrum tienen una alta participación del cliente en el proceso de desarrollo.
- En ASD y DSDM, la participación del cliente o usuario final se puede ver durante el inicio y el final de la iteración.
- Mientras que FDD utiliza informes para comunicarse con los clientes.

C. Reuniones:

La comunicación es uno de los principios básicos establecidos en el manifiesto ágil. El éxito de las metodologías ágiles depende de la comunicación efectiva entre los miembros del equipo [37]:

- Las reuniones son de carácter informal y no se conserva documentación.
- Debido al uso de la técnica de programación en pares, el éxito de XP depende en gran medida de la comunicación.
- FDD y DSDM se basan en informes y documentación para la comunicación.
- Las reuniones cara a cara se utilizan para la comunicación en ASD.

D. Tamaño y complejidad del proyecto:

Cada metodología es adecuada para un tipo particular de proyecto:

- XP y ASD son usualmente preferidos para proyectos pequeños y menos complejos.
- XP es adecuado para proyectos donde hay un cambio constante en la especificación del producto.
- Scrum, FDD y DSDM se puede aplicar a cualquier tamaño de proyecto.

En la Tabla II y Tabla III se comparan las metodologías ágiles revisadas anteriormente con respecto a diferentes parámetros [36]:

TABLA II
COMPARACIÓN DE METODOLOGÍAS ÁGILES

Características	XP	SCRUM
Enfoque	Iterativo, incremental	Iterativo, incremental
Periodo del ciclo de iteración	1 – 6 semanas	2 – 4 semanas
Tamaño adecuado del proyecto y complejidad	Proyecto pequeño y sencillo	Para proyectos grandes y complejos
Involucramiento del usuario	Activamente involucrado	A través del propietario del producto
Documentación	Documentación básica	Documentación básica
Principales prácticas	Simplicidad, Programación en pares.	Reuniones de SCRUM
Desarrollo de características concurrentes	Posible	Posible

Nota: La tabla muestra una comparación entre las metodologías XP y SCRUM (Metodologías ágiles)

TABLA III
COMPARACIÓN DE METODOLOGÍAS ÁGILES

Características	DSDM	FDD	ASD
Enfoque	Iterativo	Iterativo	Iterativo, incremental

Periodo del ciclo de iteración	En 20% por ciento del tiempo total 80% de producto.	2 días - 2 semanas	4 – 8 semanas
Tamaño adecuado del proyecto y complejidad	Todo tipo de proyectos	Proyectos a gran escala	Proyectos más pequeños y sencillos
Involucramiento del usuario	A través de lanzamientos frecuentes	A través de informes	A través de lanzamientos frecuentes
Documentación	Más que XP y SCRUM	Más alto entre todos	Documentación básica
Principales prácticas	Time boxing, MoSCoW, Prototipado	Modelado de objetos, desarrollo por característica, uso del diagrama UML	Time boxing, Risk Driven, basado en características
Desarrollo de características concurrentes	Posible	Posible	Posible

Nota: La tabla muestra una comparación entre las metodologías DSDM, FDD y ASD (Metodologías ágiles)

En la Figura 7 se puede evidenciar los resultados del XII Informe Anual del Estado de las Metodologías Ágiles del 2017, en donde se manifiesta que Scrum es la metodología ágil preferida de las organizaciones a nivel mundial:

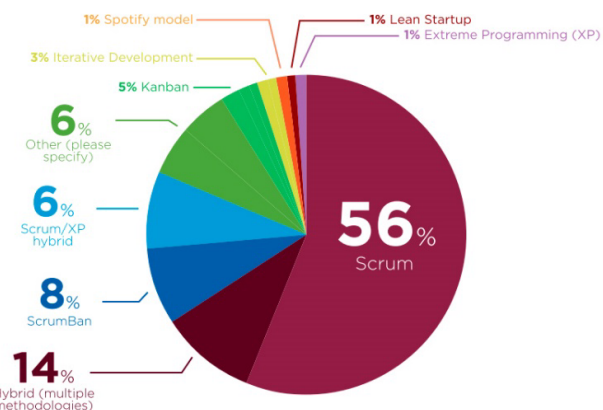


Fig. 7. Métodos y prácticas ágiles [12].

IV. CONCLUSIONES

- Las metodologías ágiles están ganando popularidad y ahora son preferidas a las metodologías de desarrollo de software tradicionales que tienen varios inconvenientes, como la incapacidad para hacer frente a los requisitos de usuario en constante cambio y exceder el tiempo y presupuesto asignados. Con los modelos de desarrollo de software tradicionales, los requisitos del producto deben especificarse claramente de antemano. Teniendo en cuenta el entorno empresarial actual, es importante que la metodología de

desarrollo utilizada se adapte fácilmente a los requisitos cambiantes del usuario final.

- Las metodologías de desarrollo de software ágil manejan los requisitos cambiantes de los clientes a través de un enfoque iterativo e incremental. Tiene una iteración de desarrollo más corta con cada iteración o incremento seguido de pruebas y análisis de riesgos que resultan en un desarrollo y entrega más rápidos de un producto de calidad.
- La selección de metodologías ágiles apropiadas es importante para maximizar la probabilidad de entrega de un producto de alta calidad que cumpla con los requisitos del usuario final. La comparación realizada en esta investigación se puede utilizar para decidir qué metodología se puede adaptar a un proyecto en particular.

REFERENCIAS

- [1] J. Mera and X. Quiñónez-Ku, "Análisis de los procesos de desarrollo de software en el departamento de TICs de la Pontificia Universidad Católica del Ecuador Sede Esmeraldas," Pontificia Universidad Católica del Ecuador Sede Esmeraldas, 2015.
- [2] O. Tinoco, P. Rosales, and J. Salas, "Criterios de selección de metodologías de desarrollo de software," *Ind. Data*, vol. 13, no. 2, p. 070, 2014.
- [3] Y. Amaya, "Metodologías ágiles en el desarrollo de aplicaciones para dispositivos móviles," *Rev. Technol. J. Technol.*, vol. 12 número, pp. 111–124, 2013.
- [4] D. A. Godoy, E. A. Belloni, H. Kotynski, H. dos Santos, and E. O. Sosa, "Simulado Proyectos de Desarrollo de Software Administrados con Scrum," *XVI Work. Investig. en Ciencias la Comput.*, pp. 485–489, 2014.
- [5] A. Navarro, J. Fernández, and J. Morales, "Revisión de metodologías ágiles para el desarrollo de software A review of agile methodologies for software development," *Univ. Icesi*, vol. 11 No. 2, pp. 30–39, 2013.
- [6] M. STOICA, M. MIRCEA, and B. GHILIC-MICU, "Software Development: Agile vs. Traditional," *Inform. Econ.*, vol. 17, no. 4/2013, pp. 64–76, 2013.
- [7] J. Canós, P. Letelier, and C. Penadés, "Metodologías Ágiles en el Desarrollo de Software," in *Métodologías Ágiles en el Desarrollo de Software*, Alicante: Universidad Politécnica de Valencia, 2003, pp. 1–8.
- [8] S. Overhage and S. Schlauderer, "Investigating the long-term acceptance of agile methodologies: An empirical study of developer perceptions in Scrum projects," in *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2012, pp. 5452–5461.
- [9] A. Moniruzzaman and S. Hossain, "Comparative Study on Agile software development methodologies," *arXiv Prepr. arXiv1307.3356*, vol. V, no. 3, pp. 37–56, 2013.
- [10] G. S. Matharu, A. Mishra, H. Singh, and P. Upadhyay, "Empirical Study of Agile Software Development Methodologies," *ACM SIGSOFT Softw. Eng. Notes*, vol. 40, no. 1, pp. 1–6, 2015.
- [11] A. Menzinsky, G. López, and J. Palacio, *Scrum Manager: Guía de formación*. Scrum Manager, 2016.
- [12] VersionOne, "12th Annual State of Agile Report," 2017.
- [13] A. Rosado, A. Quintero, and C. Meneses, "Desarrollo Ágil De Software Aplicando Programación Extrema," *Ingenio*, vol. 5, no. 1, pp. 2011–642, 2012.
- [14] P. Suardiyana, A. Yuliawati, and P. Mursanto, "Industrial Extreme Programming practice's implementation in rational unified process on agile development theme," in *2012 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, 2012, pp. 137–142.
- [15] R. Pressman, *Ingeniería.del.Software.-.Roger.Pressman.6th.Ed.McGraw-Hill.pdf*. .
- [16] Y. Yong and B. Zhou, "Evaluating extreme programming effect through system dynamics modeling," in *Proceedings - 2009 International Conference on Computational Intelligence and Software Engineering, CiSe 2009*, 2009.
- [17] I. Sommerville, *INGENIERÍA DE SOFTWARE*, NOVENA EDI. México: PEARSON, 2011.
- [18] B. Xu, "Towards high quality software development with extreme programming methodology: Practices from real software projects," in *Proceedings - International Conference on Management and Service Science, MASS 2009*, 2009.
- [19] V. Temitayo, A. Badru, and N. Ajayi, "Adopting Scrum as an Agile Approach in Distributed Software Development: A Review of Literature," *Ieee*, vol. 1, no. 2, pp. 1–5, 2017.
- [20] K. Kaur, A. Jajoo, and Manisha, "Applying agile methodologies in industry projects: Benefits and challenges," in *Proceedings - 1st International Conference on Computing, Communication, Control and Automation, ICCUBEA 2015*, 2015, pp. 832–836.
- [21] B. L. Romano and A. D. Da Silva, "Project management using the scrum agile method: A case study within a small enterprise," in *Proceedings - 12th International Conference on Information Technology: New Generations, ITNG 2015*, 2015, pp. 774–776.
- [22] A. Srivastava, S. Bhardwaj, and S. Saraswat, "SCRUM model for agile methodology," in *Proceeding - IEEE International Conference on Computing, Communication and Automation, ICCCA 2017*, 2017, vol. 2017–Janua, pp. 864–869.
- [23] A. Ahmed, S. Ahmad, N. Ehsan, E. Mirza, and S. Z. Sarwar, "Agile software development: Impact on productivity and quality," in *2010 IEEE International Conference on Management of Innovation & Technology*, 2010, pp. 287–291.
- [24] M. Mahalakshmi and M. Sundararajan, "Tracking the student's performance in Web-based education using Scrum methodology," in *Proceedings of the International Conference on Computing and Communications Technologies, ICCCT 2015*, 2015, pp. 379–382.
- [25] M. García, "Estudio comparativo entre las metodologías ágiles y las metodologías tradicionales para la gestión de proyectos software," 2015.
- [26] A. A. Abdelaziz, Y. El-Tahir, and R. Osman, "Adaptive Software Development for developing safety critical software," in *Proceedings - 2015 International Conference on Computing, Control, Networking, Electronics and Embedded Systems Engineering, ICCNEEE 2015*, 2016, pp. 41–46.
- [27] A. F. Chowdhury and M. N. Huda, "Comparison between adaptive software development and feature driven development," in *Proceedings of 2011 International Conference on Computer Science and Network Technology, ICCSNT 2011*, 2011, vol. 1, pp. 363–367.
- [28] V. P. Doshi and V. Patil, "Competitor driven development: Hybrid of extreme programming and feature driven reuse development," in *1st International Conference on Emerging Trends in Engineering, Technology and Science, ICETETS 2016 - Proceedings*, 2016, pp. 1–6.
- [29] N. R. Mead, V. Viswanathan, and D. Padmanabhan, "Incorporating security requirements engineering into the dynamic systems development method," in *Proceedings - International Computer Software and Applications Conference*, 2008, pp. 949–954.
- [30] J. Jabeen *et al.*, "Incorporating artificial intelligence technique into DSDM," in *Asia-Pacific World Congress on Computer Science and Engineering, APWC on CSE 2014*, 2014.
- [31] Priyanka and P. Kantha, "A Comprehensive Study of Traditional and AGILE Software Development Methodologies," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 6, no. 11, pp. 128–138, 2016.
- [32] M. M. Kirmani, "Agile methods for mobile application development: A

- comparative analysis," *Int. J. Adv. Res. Comput. Sci.*, vol. 8, no. 5, pp. 1200–1205, 2017.
- [33] A. Mohammad, A. Khaled, A. Nidal, and T. Ahmed, "A Comparative Study of Agile Methods: XP versus SCRUM," *Int. J. Comput. Sci. Softw. Eng.*, vol. 4, no. 5, pp. 126–129, 2015.
- [34] R. P. Pawar, "A Comparative study of Agile Software Development Methodology and traditional waterfall model," *IOSR J. Comput. Eng.*, pp. 1–8, 2015.
- [35] W. C. D. S. Carvalho, P. F. Rosa, M. D. S. Soares, M. A. T. Da Cunha, L. C. Buiatte, and M. A. T. Da Cunha, "A comparative analysis of the agile and traditional software development processes productivity," in *Proceedings - International Conference of the Chilean Computer Science Society, SCCC*, 2012, pp. 74–82.
- [36] K. Hiwarkar, A. Doshi, R. Chinta, and R. Manjula, "Comparative Analysis of Agile Software Development Methodologies-A Review," *J. Eng. Res. Appl.*, vol. 6, no. 3, pp. 80–85, 2016.
- [37] A. B. M. Moniruzzaman and S. Akhter, "Comparative study on software development methodologies.," *Database Syst. J.*, vol. 5, no. 3, pp. 37–56, 2014.

VISUAL THINKING + E-LEARNING: Relatoría gráfica en el proceso de investigación formativa

Melba-Cristina Marmolejo-Cueva, Nuria Rey-Somoza

Pontificia Universidad Católica del Ecuador-Sede Esmeraldas- Ecuador, Escuela de Ingeniería en Sistemas,
melba.marmolejo@pucese.edu.ec

Pontificia Universidad Católica del Ecuador-Sede Esmeraldas- Ecuador, Escuela de Diseño Gráfico,
nuria.rey@pucese.edu.ec

Resumen.- El aporte de las Tecnologías de la Información y Comunicación (TIC) a la comunicación y al proceso de aprendizaje es una realidad que origina el interés de investigadores en diferentes áreas, incluida la educativa y aunque es evidente la aparición de posturas que vinculan a las tecnologías con la educación, también es cierto el poco desarrollo de estudios que relacionen herramientas de comunicación visual con los espacios virtuales de aprendizaje.

La convergencia entre el Pensamiento visual o *Visual Thinking* y los espacios de aprendizaje es la propuesta que se plantea en este trabajo como una manera de garantizar la comunicación como punto de inicio dentro del proceso de aprendizaje.

Este artículo presenta la implementación de la *Relatoría Gráfica* como estrategia de síntesis del contenido de clase impartido en la escuela de ingeniería en sistemas, de cuya actividad de recogen ideas que aportan a las competencias de los estudiantes más allá de las áreas de especialidad, como el trabajo multidisciplinario, la capacidad de adaptación y el aporte de las representaciones visuales en el proceso de aprendizaje.

Palabras clave: E-learning, Visual Thinking, Relatoría Gráfica

Abstract. - The contribution of Information and Communication Technologies (ICT) to communication and the learning process is a reality that gives rise to the interest of researchers in different areas, including education and although the emergence of positions that link technologies is evident With education, it is also true that there is little development of studies that link visual communication tools with virtual learning spaces.

The convergence between Visual Thinking or Visual Thinking and learning spaces is the proposal proposed in this work as a way to guarantee communication as a starting point in the learning

process.

This article presents the implementation of the *Rapporteurship Graph* as a strategy of synthesis of the content of class taught in the school of systems engineering, from whose activity they gather ideas that contribute to the competences of students beyond the areas of specialty, such as multidisciplinary work, the ability to adapt and the contribution of visual representations in the learning process.

Palabras clave: E-learning, Visual Thinking, Graphic Rapporteurship

I. INTRODUCCIÓN

La constante búsqueda del ser humano para comunicarse de mejor manera ha promovido la creación de técnicas y herramientas que pretenden incidir en los mecanismos de intercambio de ideas dentro de la sociedad, más allá del código y su canal, para que el mensaje se perciba con claridad y perdure en el tiempo.

En cada etapa dentro de la evolución humana, se han presentado fenómenos que modifican los valores culturales, en este marco, las representaciones visuales del pensamiento son el reflejo de estos cambios, dando lugar a una cultura visual y con el surgimiento de las tecnologías sólo se ha fomentado en mayor grado el uso de representaciones gráficas para exponer ideas.

Se han presentado propuestas metodológicas, estrategias y hasta herramientas que, bajo una fundamentación teórica sólida, basada en teorías de psicología y percepción permiten crear composiciones visuales que sirven de nexo entre emisor, mensaje y receptor, sin embargo, la aplicación estratégica de estas metodologías en el aula de clase no es frecuente en algunos contextos limitados por las deficiencias en equipos o servicios tecnológicos, diezmando su potencialidad.

II. DE LA IDEA INICIAL AL VISUAL THINKING

La capacidad de pensar de manera racional es una de las características que distingue a los humanos de otras especies. La

idea que parte de la simplicidad puede evolucionar hacia un conjunto de pensamientos que al expandirse lógicamente y coherentemente a través de otros enfoques (vinculantes o divergentes), produce conceptos cada vez más profundos y complejos a medida que se desarrollan.

Las primeras civilizaciones presentan manifestaciones de comunicación relativamente claras para cada época, se presentan composiciones bien definidas y estéticas, como es el caso de las pinturas rupestres, los jeroglíficos o los vitrales. Las ideas expuestas en dicho contexto reflejan la forma de vida, así como sus principales preocupaciones, dotando de claridad al mensaje captado claramente inclusive en la actualidad.

En sus inicios las representaciones carecían de las complejidades actuales, dado el modelo básico de comunicación estructurado por emisor, medio, mensaje y receptor, incrementando la capacidad difusora del mensaje en contraposición a conceptos contemporáneos que, por su naturaleza compleja, presentan deficiencias para ser entendidos por los receptores. Este fenómeno guarda estrecha relación con el tipo de código, así como el canal utilizado para la difusión del mensaje.

Es evidente que el pensamiento evoluciona paralelamente con la sociedad y las ideas construidas en el pasado no presentan el mismo nivel de complejidad de los mensajes actuales, además de que el modelo de comunicación es diferente desde la perspectiva de la capacidad de intervención del receptor (ahora llamado usuario). Este componente dinamizador del medio-mensaje se atribuye a las nuevas tecnologías, transformando al receptor de ente pasivo a gestor activo de la información.

Cabe recalcar, que el modelo de comunicación vigente, de carácter dinámico multidireccional con capacidad de acceso a grandes volúmenes de datos y en tiempo récord, dista mucho del modelo comunicativo inicial, de carácter estático, información medida e intercambio paulatino.

Las necesidades de comunicación siguen transformándose y en este punto surgen herramientas como el *Visual Thinking* o Pensamiento Visual, cuyo propósito es simplificar conceptos complejos mediante el uso de imágenes para facilitar la comprensión del mensaje.

Tal como se menciona [1] el *Visual Thinking* es una técnica de comunicación eficaz mediante el uso de imágenes. Teóricamente fue presentada por el psicólogo y filósofo alemán Rudolf Arnheim en el año 1969 y [2] declara que la inteligencia no es posible sin la percepción, debido a su reciprocidad con el pensamiento; sin embargo el reconcomiendo y difusión actual como técnica se atribuye a Dan

Roam, quien propone un proceso compuesto por 4 pasos: mirar, ver, imaginar y mostrar. En conjunto permiten exponer el mensaje, indagar sobre él, observar sus relaciones internas y comprenderlo más allá del vistazo inicial. Además, por su efectividad, esta técnica se ha implantado como la herramienta de comunicación con mayor aceptación dentro de procesos creativos para compañías reconocidas como exitosas como Apple, Google o Ebay.

Más allá de la postura herramental, se define al *Visual Thinking* como la capacidad innata del individuo durante el descubrimiento de sus propias ideas que al ser procesadas se convierten en conceptos que requieren de elementos visuales para ser compartidos de manera rápida e intuitiva [3]. En este punto, la neuroplasticidad junto a otros mecanismos propios del cerebro tienen un papel fundamental porque permiten procesar las señales visuales, considerando que el principal órgano de visión es el cerebro y no los ojos, ya que éste busca constantemente patrones para organizar las percepciones, aplicando las leyes gestálticas. La percepción de un elemento desconocido provoca un estímulo visual mayor comparado con otro objeto conocido. Esta interacción intelecto-estímulo es parte del proceso de comprensión y asimilación de conocimiento.

Investigaciones que afirman que las imágenes mentales son el primer paso antes de comunicar una idea obtenida tras la acción de pensar [4]. El *Visual Thinking* activa las células neuronales, provocando la expresión de la razón y emociones. Este estudio también señala que neurológicamente los dibujos conectan la intelectualidad, creatividad, compendios psicológicos y otras características para lograr:

- Unión mente - emoción para convertir lo imaginario (intangible) en objeto real (tangible) tras la visualización.
- El nivel de representación realista de la idea-concepto no afecta su percepción.
- Visión perceptiva compuesta por ojo + mente + espacio de representación es posible porque el cerebro puede reconocer las representaciones visuales abstraídas e imperfectas como una realidad.
- La acción de convertir en realidad lo observado a través de su representación visual es una premisa para el aprendizaje.

Todas estas aportaciones, a largo plazo, buscan resolver la forma de comprender el funcionamiento del cerebro enmarcado en los fundamentos de la neurociencia con el propósito de identificar las herramientas y técnicas efectivas del *Visual Thinking*.

III. HERRAMIENTAS TECNOLÓGICAS EN EL PROCESO DE APRENDIZAJE

Para nadie es ajeno el uso cotidiano y diversificado de las herramientas tecnológicas dentro de la sociedad moderna, a tal punto que las plataformas digitales se han convertido, en algunos casos, en espacios ineludibles para la ejecución de actividades

como la gestión de datos e incluso de información destinada a la ciencia y el conocimiento.

Básicamente, en esta nueva línea de trabajo orientado a la modernización de las disciplinas consideradas como clásicas, el proceso educativo está mutando hacia una formación ampliada de los estudiantes, otorgando las competencias adicionales para su desempeño profesional, en este punto, se afirma que las herramientas tecnológicas pueden ser un aporte sustancial para dicho fin al promover capacidades cognitivas como el razonamiento, síntesis y selección para un aprendizaje basado en la experiencia [5].

A partir de la adopción de las plataformas digitales en los espacios educativos, surge un modelo educativo denominado como *e-learning*, que es una aplicación web que integra varias utilidades destinadas al proceso de enseñanza-aprendizaje en la Internet [6].

Si bien la esencia del *e-learning* es la interacción síncrona o asíncrona entre docente y alumno, durante el proceso de formación, se presenta como principal atributo la capacidad autónoma del estudiante para acceder a sus clases según sus necesidades y ritmo de aprendizaje propio. Dentro de este modelo, también se reconoce al *b-learning* como la combinación entre el aprendizaje autónomo por medio de plataformas virtuales y la experiencia de clases presenciales.

Aunque la aplicación tanto del *e-learning* como del *b-learning* en el área educativa no es algo nuevo, lo cierto es que la construcción del mensaje para estos espacios de aprendizaje se ha desligado de los modelos de comunicación que constituyen un referente fundamental para asegurar la claridad del mensaje a difundir en dichos espacios.

Las herramientas tecnológicas siguen mejorando sus características técnicas, no obstante, los espacios de aprendizaje requieren de otros componentes con igual relevancia y que deberían ser priorizados en las primeras fases de desarrollo de las aplicaciones digitales. Es aquí donde se sugiere al *Visual Thinking* como el instrumento idóneo con capacidad de potencializar el proceso educativo en los espacios digitales.

IV. CONVERGENCIA ENTRE EL VISUAL THINKING Y EL E-LEARNING

Si bien es evidente el aporte de las plataformas tecnológicas para el campo educativo, también se reconoce su complejidad dentro del proceso de implementación debido a acciones de gestión realizadas por personal no informático (por

ejemplo, personal administrativo y docente de una institución educativa) [6].

En este punto es donde el *Visual Thinking* se hace presente como mecanismo de innovación para mejorar sistemas o ideas. Al exponer que los elementos gráficos facilitan la resolución de problemas por ser una actividad constructiva que tiende a la creación de ideas innovadoras, además de que su implementación garantiza la memorabilidad del mensaje y se confirma su efectividad como transmisor de mensajes destinados al ámbito educativo.

Entre los recursos del *Visual Thinking* acordes al área educativa están:

- Diagramas y mapas mentales. – Se presentan como representaciones lógico-analíticas a través de la relación entre ideas clave que se distribuyen en una composición con un concepto central del que convergen las ideas consecuentes.
- Storyboard. – Definido como una secuencia gráfica que representa una historia. Se origina como herramienta narrativa en la industria cinematográfica, sin embargo, también es utilizada para la presentación de ideas relacionadas con el área comercial o educativa.
- Procesos lógicos. – Se utilizan para representar las relaciones secuenciales entre sus componentes, pudiendo visualizar aspectos como las emociones que vive el usuario en función del punto del proceso en el que se encuentre.
- Modelo. – Se define como una representación de tipo exploratorio y emocional, generalmente ejecutado con simplicidad como un boceto a mano alzada [7].

El entorno en que se desenvuelve la sociedad digital, en constante cambio y permanente conexión, está sobre estimulado debido a la disposición de todo tipo de información presentada de en formatos heterogéneos, produciendo contenidos cada vez más complejos y no siempre fiables. Bajo este argumento, el hiperrealismo y la proliferación de imágenes incoherente con la realidad del entorno disminuyen el valor significativo de los contenidos expuestos en la Internet y el desbordamiento de información constituye un reto para la labor educativa [8].

Se presenta la relevancia de las representaciones visuales en el modelo cognitivo para formar alternativas de mediación en el aprendizaje enmarcado en la tecnología y gracias a la virtualidad, su implementación en espacios lúdicos es un hecho que incrementa el campo de acción formativa [9] y se coincide en la importancia de relacionar el pensamiento visual con la psicología cognitiva además de complementarse con otras teorías como: el color, la semiótica, tipos de aprendizaje [4].

La capacidad de integración entre los componentes del pensamiento visual, los modelos cognitivos y espacios tecnológicos destinados a la educación a través del e-learning, constituyen la base para proponer modelos de comunicación y educación innovadores, pero sobre todo adaptadas al nuevo paradigma de interacción de información originado por las TIC.

El potencial de las tecnologías es relativo y depende de las necesidades de formación, así como el nivel de aporte en el proceso formativo, en este sentido la alfabetización digital es un requerimiento para el desarrollo eficiente del modelo educativo planteado.

Por otro lado, se aclara que el diseño de una sola estrategia en las e-learning, aunque se potencialicen con el pensamiento visual no garantizan el aprendizaje significativo si durante la planeación del recurso no se considera el estilo de aprendizaje, esto no significa que las propuestas desarrolladas a partir de la generalidad no sean válidas, pero para generar modelos efectivos es necesario indagar en disciplinas como la percepción, representación, cognición y aprendizaje. De su correcta configuración depende la proyección hacia una cultura visual digital, que permita convertir a la representación gráfica en información y posterior conocimiento [10].

V. CASO PRÁCTICO: RELATORÍA GRÁFICA CLASES DE INGENIERÍA EN SISTEMAS PUCESE 2018

En los referentes teóricos anteriormente tratados se mencionan algunos productos derivados de la técnica *Visual Thinking*, como es el caso de la *Relatoría Gráfica*, procedimiento que consiste en recoger de manera gráfica un suceso o evento. Una de las figuras que destaca en este método es Carla Boserman [7] utilizándolo como proceso y método de investigación desde la visualidad.

De su aplicación en los espacios digitales han resultado productos de comunicación con valor lúdico fácilmente asociados al e-learning y b-learning.

la *Relatoría Gráfica* o el *relatograma* se concibe como una forma de explorar el contexto y otorgarle visibilidad a partir de la relación de ideas e imágenes. Por lo tanto, sirve como método y resultado de investigación al priorizar la manera de obtener datos, así como la representación final de los mismos.

Esta propuesta fue desarrollada por estudiantes de la PUCSE con la escuela de Diseño Gráfico a partir de una experiencia de clase con estudiantes de la escuela de Ingeniería en Sistemas. La elaboración de estos productos gráficos formó parte del trabajo dedicado al tiempo de investigación formativa, mediante la estrategia de investigación acción, compuesta por las 5 fases:

- a) **Diagnóstico y exploración.** - observando ejemplos de *Visual Thinking* y *Relatoría Gráfica*)

- b) **Planificación.** - Selección de materias, cronograma de trabajo, objetivos)

- c) **Metodología.** - Aproximación a las técnicas de *Visual Thinking* en procesos de investigación)

- d) **Análisis e interpretación.** – Adentrarse en el contexto de acción y generar los productos durante el proceso

e) Finalización y socialización de los productos gráficos

Los estudiantes están experimentando con métodos de investigación más próximos a su campo de conocimiento, sin embargo, la interacción con otros espacios ajenos a su especialidad, se logra reconocer otros enfoques (cuantitativo y cualitativo) de investigación. Aunque, en la construcción fue necesario insertarse en un contexto diferente a su espacio de aprendizaje habitual, comparte aspectos relacionados con el valor y análisis cualitativo.

Desde la materia de Taller de Diseño VII, los estudiantes de Diseño Gráfico tienen entre sus competencias la creación de representaciones infográficas bi y tridimensionales fundamentados con los conocimientos adquiridos de comunicación, composición y síntesis de la imagen. La última unidad de la materia, dedicada a modos alternativos de construcción de infografías, se centra en explorar diferentes técnicas para sus propuestas de comunicación y surge el *Visual Thinking* como estrategia de trabajo.

Para desarrollar esta actividad, el grupo de estudiantes (16 estudiantes en total) se dividió en 4 grupos, distribuidos a diferentes clases de la carrera de Sistemas. De esta manera, se convierten en partícipes activos de la clase como observadores activos y generan evidencia del contenido impartido a través de la *Relatoría Gráfica*.

Como resultado de esta intervención se presentan los siguientes relatogramas:



Fig. 1 Relatograma. Diseño: Yamilka Tomalá. Clase de TIC. Escuela de Ingeniería en Sistemas, PUCSE 2018

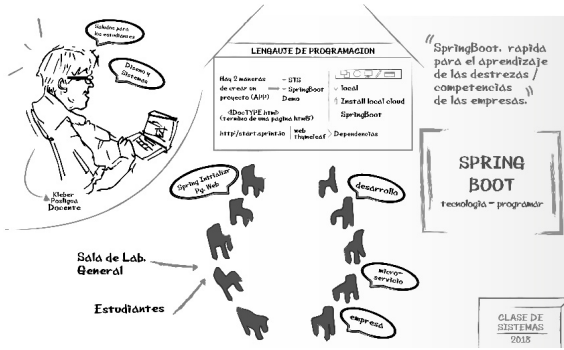


Fig. 2 Relatograma. Diseño: Stefania Angulo. Clase de Programación. Escuela de Ingeniería en Sistemas, PUCES 2018



Fig. 3 Relatograma. Diseño: Jasmine Reyna. Clase de Programación. Escuela de Ingeniería en Sistemas, PUCES 2018

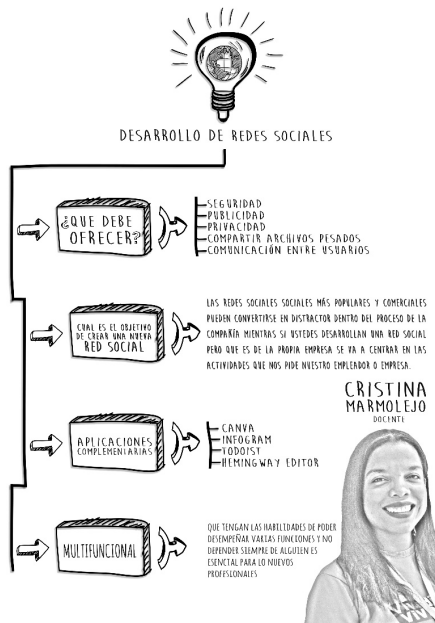


Fig. 4 Relatograma. Diseño: Daniel Céleri. Clase de TIC. Escuela de Ingeniería en Sistemas, PUCES 2018

VI. APORTES FINALES

Esta experiencia de aprendizaje mediante la integración del *Visual Thinking* como táctica para el *e-learning* busca presentar alternativas innovadoras y consistentes con los paradigmas actuales de comunicación destinada a la educación en diversos niveles, no obstante, este artículo es sólo una aproximación hacia las posibilidades ofrecidas por los métodos y técnicas antes expuestos.

A continuación, se presentan algunas reflexiones recogidas después del caso aplicado:

- Estudiantes de distintas carreras comparten espacios, profesores y conocimientos. Aunque al principio de la actividad se percibió algo de resistencia por parte de los sujetos de estudio observados en clase, Posteriormente se logró compartir dinámicas de aprendizaje comunes para todos los participando logrando identificación entre todos.
- Se pone de manifiesto el carácter transversal del diseño gráfico, así como de otras disciplinas (ilustración, fotografía, dibujo al natural) que pueden cooperar en diferentes niveles y ramas de conocimiento en favor de proyectos comunes.
- En la práctica los estudiantes se empoderan de su rol como diseñadores del pensamiento visual con orientación al aprendizaje en los nuevos espacios digitales. Este proceso de transición les resulta natural considerando que pertenecen a una generación que se identifica con los nuevos modelos de comunicación soportados por TIC, y este fenómeno trasciende de la especialidad.

VII. RESULTADOS DE LA EXPERIENCIA

- Ruptura de rutina: los estudiantes de diseño salen de su aula y laboratorio habitual para integrarse en otros espacios y contextos de la universidad.
- Aproximación a otros contextos con motivo de procesos de investigación, dejando de actuar únicamente como estudiantes de diseño para asumir roles de diseñador profesional y/o diseñador investigador
- Capacidad de integrar en los equipos de trabajo a estudiantes de especialidades diferentes como una forma de acercarlos al trabajo cooperativo y multidisciplinario que es una competencia altamente requerida por los empleadores.

- Identificación de las necesidades comunicativas para proponer soluciones con alcances lúdicos por medio de la aplicación del Visual Thinking. Fue necesario introducirse en el espacio de trabajo (clases de ingeniería en sistemas) para indagar, conocer, entender y construir las *Relatorias Gráficas*.
- Legitimación de procesos y productos de investigación artística, con viabilidad para emplearse en disciplinas de carácter técnico como es el caso de la ingeniería en sistemas, en dónde la representación gráfica ya no se concibe como un elemento intrascendente para convertirse en un complemento relevante dentro del proceso de comunicación-interacción

Thinking, el empleo de recursos gráficos para optimizar nuestros proyectos,» de *Jornada INGEGRAF. Contribución de la ingeniería gráfica a la sociedad*, Zaragoza, 2016.

- [8] M. Jiménez Montano y J. Flores Ramos, «La educación de la mirada: percepción, pensamiento complejo y competencias visuales,» de *Conferencia Interacciones*, 2016.
- [9] H. Durán Macedo, «Visualidad y conocimiento: La visualidad como forma de conocimiento, del ágora a los nuevos escenarios virtuales,» *RU-TIC. Repositorio Universitario UNAM*, vol. 16, nº 10, 1 10 2015.
- [10] I. Jaramillo Urritia, «El mapa conceptual como estructura de representación de conocimiento en cursos virtuales y su impacto en el aprendizaje visual de estudiantes adultos,» *Innovación Educativa*, vol. 15, nº 68, mayo/agosto 2015.
- [11] C. Boserman, «Entre grafos y bits,» *Revista de comunicación*, nº 6, pp. 8-23, 2014.

REFERENCIAS

- [1] Y. Iglesias, «<https://designthinking.gal/el-visual-thinking/>,» 9 febrero 2016. [En línea]. Available: <https://designthinking.gal>. [Último acceso: 16 agosto 2018].
- [2] S. Sánchez Mancha, «www.guindo.com,» 22 junio 2012. [En línea]. Available: www.guindo.com/blog/rudolf-arnheim-y-el-pensamiento-visual/. [Último acceso: agosto 2018].
- [3] J. Cantón, «Pensamiento Visual para la creatividad y la narrativa mediante herramientas digitales,» Universidad Internacional de Andalucía, Andalucía, 2017.
- [4] N. Puñez Lazo, «El Pensamiento visual: una propuesta didáctica para pensar y crear,» *horizonte de la ciencia. revista científica*, vol. 7, nº 12, 2017.
- [5] M. S. González González, «Estrategias para trabajar la creatividad en la Educación Superior: pensamiento de diseño, aprendizaje basado en juegos y en proyectos,» *RED. Revista de Educación a Distancia*, nº 40, pp. 1-15, 29 Julio 2015.
- [6] A. Cesteros Fernández-Pampillón, «Las plataformas e-learning para la enseñanza y el aprendizaje universitario en Internet. Las plataformas de aprendizaje. Del mito a la realidad,» Universidad Complutense de Madrid, Madrid, 2009.
- [7] R. Sanz-Segura y C. Romero-Piqueras, «Visual



Pontificia Universidad
Católica del Ecuador