



Actas TICAL2019

Novena Conferencia de
Directores de Tecnología
de Información y
Comunicación en
Instituciones de
Educación Superior

“El genoma estudiantil y
la metamorfosis digital
universitaria”

Actas TICAL2019

**Novena Conferencia de Directores de
Tecnología de Información y Comunicación
en Instituciones de Educación Superior**

**“El genoma estudiantil y la
metamorfosis digital
universitaria”**

Cancun International Convention Center (Cancun ICC), 2 – 4 septiembre, 2019
Cancún, México

Comité de Programa:

- ERNESTO CHINKES
Presidente Honorario, Argentina
- LEANDRO GUIMARÃES
Director Adjunto Escuela Superior de Redes, RNP,
Brasil
- ALONSO CASTRO
Director del Centro de Informática, Universidad de
Costa Rica, Costa Rica
- LUIS A. NÚÑEZ
Físico, Universidad Industrial de Santander,
Bucaramanga - Colombia
- LUIS ALBERTO GUTIÉRREZ DÍAZ DE LEÓN
Coordinador General de Tecnologías de Información
Universidad de Guadalajara, México

Compiladora, coordinadora general de la publicación, edición y diseño:
María José López Pourailly, RedCLARA.

RedCLARA (<http://www.redclara.net>)

Fecha en que se terminó la presente edición: 08-10-2019

ISBN: 978-956-9390-12-8

Copyright de la presente edición:

Actas TICAL2019 - Novena Conferencia de Directores de Tecnología
de Información y Comunicación en Instituciones de Educación
Superior

“El genoma estudiantil y la metamorfosis digital universitaria”

por RedCLARA,

se encuentra bajo una Licencia Creative Commons

Atribución-NoComercial-SinDerivadas 3.0 Unported.

Índice

Capítulo 1, Tecnologías en la Innovación: Enseñanza	9
Desarrollo de un prototipo de juego digital educativo para las Ciencias de la Tierra	11
Ambiente de coworking y co-creación en red para fortalecer el proceso de aprendizaje e investigación en la formación académica-científica. Caso UTN	28
Jugar Web: Módulo de programación visual y motor de visualización para la creación de multimedia educativa interactiva y el desarrollo del pensamiento computacional	40
Capítulo 2, La conectividad y la infraestructura en la universidad digital	57
Modelo de clasificación de prioridad para la prestación de servicios de infraestructura tecnológica en la Universidad Estatal a Distancia	59
Dockerización de una plataforma Moodle para migrarla hacia la nube	73
Experiencias recientes en la mejora de la metodología para gestión de proyectos de infraestructura TIC en el Centro de Informática de la Universidad de Costa Rica	95
Capítulo 3, La gestión y gobierno de las TIC con miras a la transformación digital universitaria	111
Trascendencia de la gestión de los servicios de TI en la vida universitaria	113
Caso de la Universidad del Rosario: Implementación del Sistema Integral de Información	125
Estrategia de implementación para los sistemas de gestión de servicios de TI: Data Center Infrastructure Management y SCG	136
Capítulo 4, Tecnologías en la Innovación -Enseñanza e Investigación	149
Implementación de un modelo de educación virtual predictiva que evita el fracaso asociado a bajos promedios de calificación	151
UnaCloud: Ejecutando Aplicaciones HPC en las Salas de Computadores de la Universidad	165
Implementación de un Sistema de Gestión de la Investigación (CRIS) utilizando tecnologías abiertas: El caso de la Universidad Autónoma de San Luis Potosí	179
Capítulo 5, Smartcampus y tendencias – Blockchain	195
SIU-Guaraní y la Universidad Digital: Registro de información académica con Firma Digital y Blockchain	197

Uso de <i>smart contracts</i> para validación de certificados en una red <i>blockchain</i>	211
Ambiente integral para fortalecer la seguridad de la información en los sistemas de la Universidad de Guadalajara	228
Capítulo 6, Respondiendo a la universidad cambiante de la era digital mediante desarrollos rápidos y flexibles	243
Sistema integral para la implementación de una credencial universitaria	245
A teoria e a prática da gestão dos documentos digitais: o caso da Universidade Federal de Santa Maria	257
Arquitectura Basada en Microservicios para Aplicaciones en Alta Disponibilidad	269
Sistema de Recomendación en la Nube para Apoyar la Colaboración entre Investigadores	282
Capítulo 7, Tecnologías en la Innovación - Investigación	297
Experiencia de la Universidad del Rosario en la Implementación de Servicios y Plataformas para Computación de Alto Desempeño y de Apoyo a la Investigación	299
Centro de Análisis de Datos y Supercómputo (CADS) de la Universidad de Guadalajara	316
Integración de las de las plataformas (georreferenciación, recursos de investigación) para la Gestión de la de investigación: caso de la Universidad Nacional, Costa Rica	329
Capítulo 8, Logrando universidades digitales seguras	339
Automatización en la seguridad del perímetro de la red mediante tecnologías open-source	341
Establecimiento del Plan de Continuidad del Centro de Operaciones de Videoconferencia de la DGTIC UNAM	353
Combate à Fraudes em Documentos Acadêmicos: Registro, Autenticação e Preservação de Diplomas Digitais Usando <i>Blockchain</i>	368
Capítulo 9, <u>Smartcampus</u> y Tendencias - Analítica de Datos/IoT	389
Sistema de Registro de Asistencia basado en IoT y Tarjeta Universitaria Inteligente para la Universidad de Valparaíso, un ejemplo de Smartcampus	391
Extracción de datos desde una plataforma de administración de equipos que brindan servicios de internet y su procesamiento para desarrollo de un tablero de indicadores	408
La implementación de las Tecnologías de la Información y Comunicación en la	

Universidad Ecuatoriana aseguran el análisis de datos y la toma de decisiones en procesos académicos 422

Capítulo 10, Transformando los servicios hacia la mejora de la experiencia del usuario 433

Gestión del conocimiento para la atención a usuarios y soporte de servicios de TI en la Universidad Autónoma de Yucatán 435

QRPower@UNGS-Activa La transformación digital universitaria a través de soluciones simples con códigos QR 455

Tablero de control universitario basado en el modelo de acreditación institucional ecuatoriano 471

Capítulo 1

Tecnologías en la Innovación: Enseñanza

Desarrollo de un prototipo de juego digital educativo para las Ciencias de la Tierra

Edgar Ordoñez-López¹, Marisol Escorza-Reyes¹

¹ Universidad Autónoma del Carmen, Facultad de Química, Calle 56 no. 4 Esq. Av. Concordia Benito Juárez, cp. 24180 Ciudad del Carmen, Campeche, México.
en.rin-1605@hotmail.com, mescorza@pampano.unacar.mx

Resumen. En el presente trabajo se muestra el desarrollo e implementación de una metodología para la creación de juegos digitales educativos basado en teorías de enseñanza-aprendizaje que sirva como herramienta para la mejora de la comprensión de un tópico de Ciencias de la Tierra para estudiantes universitarios. La metodología se planteó con la modificación del flujo de trabajo de creación de juegos digitales comerciales y se llevó a la práctica con el tema de índices de meteorización química, que pertenece al tema de Geoquímica. El juego digital, que se encuentra en fase inicial de creación, fue desarrollado bajo la modalidad *Endless Runner*, con distintos objetivos que cambian de dificultad dependiendo de los ambientes naturales que simula el juego. Este juego digital tuvo un proceso de planteamiento de la idea, definición de mecánicas que conforman al sistema del juego y la producción de todo el sistema. La creación del juego digital tiene como plataforma objetivo el celular, que es un dispositivo que se encuentra al alcance de la gran mayoría de estudiantes universitarios.

Palabras Clave: Aprendizaje basado en juegos digitales, Geoquímica, Ciencias de la Tierra.

Eje temático: Tecnologías en la innovación educativa e investigación.

Abstract. This paper shows the development and implementation of a methodology for the creation of educational digital games, that it's based on teaching and learning theories that serve as a tool for improving the understanding of an Earth Science topic for university students. The methodology was proposed with the modification of the development workflow of commercial digital games and was carried out with the subject of chemical weathering index, which pertains to Geochemistry. The digital game, which is in the initial phase of creation, was developed under the Endless Runner mode, with different objectives that change in difficulty depending on the natural environments simulated by the game. This digital game had a process of approach first from the idea, then the definition of mechanics that make up the game system and the production of the entire system. The creation of the digital game has the cell phone as a target platform, which is a device that is available to the vast majority of university students.

1. Introducción

En la actualidad las técnicas de enseñanza-aprendizaje que mejoren la eficacia de la asimilación de información en las universidades son cada vez más necesarias como parte de la innovación educativa. El aprendizaje, cualquiera que sea la forma o contexto en el que se desarrolle, al ser un proceso complejo con distintos factores que lo afectan, tiene muchas posibilidades de ser mejorado, y actualmente existen diversas investigaciones que desarrollan nuevas técnicas de enseñanza-aprendizaje, en la mayoría de estas se enfocan en la implementación de las nuevas tecnologías en el aula de clase.

En la educación superior, cada vez es más notorio que las nuevas generaciones de estudiantes han perdido el interés en las formas tradicionales de enseñanza en las aulas de clase, esto debido a que solo son receptores de conocimiento y tienen muy poca interacción que motive su proceso de aprendizaje. Además, en el contexto actual, cada estudiante recibe estímulos de diferentes fuentes: las internas o inherentes a la clase, como son el profesor y el pizarrón o proyección, o de agentes externos, como el celular o equipos de cómputo, lo que dificulta una transferencia de información óptima.

En la Universidad Autónoma del Carmen, donde se imparte el Programa Educativo en Ingeniería Geológica, el rezago estudiantil y los altos índices de reprobación son altos, y una de las causas es la falta de motivación para asistir a clases.

Por lo anterior, es de interés para los profesores universitarios implementar las nuevas tecnologías al proceso de enseñanza-aprendizaje, como respuesta a las necesidades actuales de los estudiantes, quienes se encuentran en un mundo en constante cambio producto del acelerado avance tecnológico, por lo que se requiere de estímulos diferentes para captar su atención en el aula de clase.

Entre las técnicas más innovadoras implementadas para cambiar y mejorar los estímulos de los procesos de enseñanza-aprendizaje, y que involucran la aplicación de las nuevas tecnologías, se encuentran las metodologías de enseñanza basada en juegos digitales (Khine, 2010). Cabe agregar que se apoyan de teorías de aprendizaje, sobre la forma en que el cerebro aprende, y de la enseñanza, referente a los pasos óptimos para una adecuada transferencia de la información.

Sin embargo, como la metodología de fabricación de juegos digitales comerciales contiene objetivos específicos diferentes a los de un juego digital con propósito educativo, se modificó una metodología que utiliza el proceso estándar de fabricación de juegos digitales y conceptos de las teorías de enseñanza-aprendizaje.

Así entonces, en este trabajo se aborda el desarrollo e implementación de un prototipo de juego digital para la enseñanza de un tópico de Ciencias de la Tierra, específicamente para el área de la Geoquímica, el cual busca incrementar la eficacia de la enseñanza, así como la motivación de los estudiantes, en el tema de índices de meteorización química.

La disposición del presente trabajo es la siguiente: en la sección 1 se presenta una introducción, en la sección 2 se describen los antecedentes teóricos, posteriormente, en la sección 3, se aborda la creación de la metodología de construcción de un juego digital con enfoque educativo; en la sección 4, se describe la aplicación de la metodología con el tema de índices de meteorización química y finalmente en la sección 5 se presentan las conclusiones, limitaciones y el trabajo a futuro.

1. Antecedentes

2.1 Teorías de la enseñanza y el aprendizaje a través de juegos digitales

Desde finales del siglo XIX y principios del siglo XX, el interés por comprender la forma en la que el cerebro humano aprende dio como resultado el surgimiento de diversas Teorías del Aprendizaje y la Enseñanza (Illeris, 2018).

En la actualidad, las investigaciones relacionadas a los procesos de enseñanza-aprendizaje son numerosas, sin embargo, muchas de ellas se centran en utilizar los conceptos de las teorías existentes para generar metodologías que implementen nuevas herramientas tecnológicas. Tal es el caso de las investigaciones relacionadas al aprendizaje basado en juegos digitales (Blunt, 2007; Ashinoff, 2014; Crocco, Offenholley, & Hernandez, 2016; Qian & Clark, 2016).

Estudios recientes muestran que el aprendizaje basado en juegos digitales tiene influencia en los diferentes componentes del aprendizaje: cognitivo, afectivo, de comportamiento y sociocultural (Plass, Homer, & Kinzer, 2015) y que además, tiene el potencial de contribuir con el denominado aprendizaje para toda la vida o *lifelong learning* (Pivac, Dziabenko, & Schinnerl, 2003; El Mawas & Muntean, 2018; Oanh, 2018). De igual forma, debido al avance tecnológico y las nuevas formas de despliegue de los juegos digitales, el aprendizaje se puede llevar con más eficacia, al conectar el mundo real con el mundo virtual a través de experiencias inmersivas (M. C. Li & Tsai, 2013).

Así entonces, para alcanzar a comprender la forma en la que el cerebro aprende, a continuación se enuncian las teorías del aprendizaje que representan una mayor relevancia para el planteamiento de la metodología de creación de un juego digital con fines educativos.

Teoría del Aprendizaje por descubrimiento. Define que aprendemos al crear una hipótesis sobre cómo funciona nuestro entorno y luego experimentamos para probar esa hipótesis (Bruner, 1961). Esta teoría se centra en que el aprendizaje se adquiere por actos de descubrimiento, en donde se incluyen todas las formas de obtener conocimiento por uno mismo y solo utilizando la propia mente (Seel et al., 2017).

Teoría de la Carga cognitiva. Todos tenemos una capacidad limitada de atención en nuestra memoria de trabajo o la memoria de corto plazo, y el aprendizaje puede ser facilitado cuando ayudamos al usuario a reducir la demanda de memoria (Sweller, 1988).

Teoría de la autodeterminación. Descrita por Edward Deci, menciona que existen tres motores que son fundamentales para la motivación humana para realizar cualquier actividad a realizar, estos motores son: competencia, autonomía y relación, son necesidades psicológicamente innatas que mueven a todo ser humano (Ryan & Deci, 2000). Algo debe de estimular esas necesidades antes de que un individuo vea una actividad como motivante.

Competencia significa poder demostrar que tenemos la capacidad de lograr algo, es decir, ser reconociblemente bueno en algo. *Autonomía* es sobre la independencia y la libertad de elegir por uno mismo. *Relación* se refiere al sentimiento de estar conectado a otros y que sea de importancia para ambas partes.

Las teorías de aprendizaje consideran cómo las condiciones existentes de una persona y su entorno interactúan con el funcionamiento psicológico y fisiológico de la mente y el cuerpo humano para lograr un cambio en la visión del mundo, el comportamiento y/o las habilidades de esa persona (Becker, 2017).

En el aula de clase no se puede modificar las condiciones mentales existentes de cada alumno, pero el entorno de enseñanza si se puede cambiar. Ya que se crea a base de instrucciones, métodos y secuencias lógicas, es aquí donde las teorías de la enseñanza proveen los marcos de referencia para formular como dar una instrucción de manera efectiva. Una teoría de enseñanza ofrece una guía explícita sobre cómo ayudar mejor a las personas a aprender y desarrollarse (Reigeluth, 1999).

Teoría de Gagné y los nueve eventos. Describe una serie de eventos que propician que una lección sea efectiva (Gagné, 2008). Los nueve eventos son:

1. Ganar atención (*recepción*), al dar pistas del estilo, historia y objetivos principales.
2. Informar a los aprendices el objetivo (*expectativa*), comunicar el objetivo principal que debe coincidir con la finalidad del aprendizaje.
3. Estimulación de aprendizaje previo (*recuperación*), los retos deben estar diseñados para hacer uso del conocimiento previo.
4. Presentación de los estímulos (*percepción selectiva*), el modo de presentar el contenido al usuario.
5. Dando guía al aprendizaje (*codificación semántica*), dar guía para ayudar a los aprendices a codificar y asimilar lo que están aprendiendo.
6. Provocar mejor rendimiento (*respuesta*), los aprendices necesitan que se les dé oportunidad de practicar lo que han aprendido.
7. Dar retroalimentación (*refuerzo*), el usuario debe saber dónde se encuentra, al mismo tiempo de que conozca su puntaje o rendimiento.
8. Evaluar el desempeño (*recuperación*), al proveer un resumen una vez que el trayecto o nivel se ha completado.
9. Mejorar la retención y la transferencia (*generalización*), al tener la posibilidad de transferir conocimiento de una actividad a otra.

Primeros principios de Merrill. Consiste en una serie de elementos fundamentales para toda instrucción, que se han sintetizado de otras teorías y modelos desarrollados por Merrill (David Merrill, 2009). Estos elementos son:

- *El problema.* El aprendizaje se facilita cuando el aprendiz está comprometido en resolver un problema real en el mundo.
- *Activación.* El aprendizaje se facilita cuando el conocimiento existente es activado y es el cimiento de nuevo conocimiento.
- *Demostración.* El aprendizaje se facilita cuando el nuevo conocimiento se le ha reconocido al usuario. Al mostrar los resultados y no solo contarles que lo lograron.
- *Aplicación.* El aprendizaje se facilita cuando el nuevo conocimiento es aplicado por el aprendiz en cuestión.
- *Integración.* El aprendizaje se facilita cuando el nuevo conocimiento es integrado en el mundo del aprendiz. Si el aprendizaje es totalmente inerte, se desvanece.

2.2 Conceptos de Ciencias de la Tierra y Geoquímica

El objetivo de la Geoquímica, como disciplina de las Ciencias de la Tierra, es el estudio de la composición de la Tierra y sus partes, así como de la distribución y movilidad de los elementos químicos (Misra, 2012).

En la corteza terrestre existen diferentes tipos de rocas, las cuales se diferencian entre sí por tener composiciones químicas y mineralógicas distintas. No obstante, una roca no permanece igual durante su ciclo sobre la superficie terrestre, y más bien, cambiará su forma, tamaño y composición química conforme se encuentre bajo la influencia de los agentes ambientales como el agua y el viento; a estas transformaciones se les denomina meteorización.

La meteorización puede definirse de manera formal como un proceso de la naturaleza que se produce cuando la roca es fragmentada mecánicamente (desintegrada) o alterada químicamente (descompuesta). La meteorización mecánica se lleva a cabo por fuerzas físicas que rompen la roca en fragmentos más pequeños sin modificar la composición química de la roca, mientras que la meteorización química implica un cambio en la composición química de los componentes de la roca (Tarbuck & Lutgens, 2005).

Durante la meteorización química ocurren procesos complejos que modifican los minerales de las rocas y sus estructuras internas. Dichos procesos convierten los constituyentes de los minerales originales en otros nuevos o los liberan al ambiente circundante. Los principales mecanismos de meteorización química son:

Disolución. Las rocas se disuelven, produciendo la solubilización de iones, cuando son afectadas por aguas meteóricas cargadas de distintos ácidos, sobre todo el carbónico de la atmósfera y el húmico del suelo.

Oxidación. Se produce cuando un elemento pierde electrones durante la ocurrencia de reacciones químicas.

Hidrólisis. Consiste en la reacción que produce la ruptura de cualquier sustancia por la interacción con la molécula del agua. En la naturaleza, el agua contiene normalmente otras sustancias que incrementan la concentración de iones de hidrógeno, lo cual acelera la reacción de hidrólisis; la presencia de una cantidad incluso pequeña de ácido aumenta de una manera notable la fuerza corrosiva del agua (Shick, Hoffmann, Lamb, & Hole, 2015).

Los mecanismos de meteorización química producen movilidad de elementos y, aunque es un fenómeno que ocurre en diferentes contextos geológicos, en el análisis de ambientes geológicos superficiales, donde la movilidad de los elementos está controlada principalmente por el agua, el uso de los índices de meteorización química es habitual (Y. S. Li, Zhao, & Yang, 2010; Roy et al., 2015).

Los índices de meteorización química son estimaciones de la pérdida de elementos móviles en un material geológico por la acción de diversos factores ambientales y el tiempo. Los principales elementos móviles que se analizan en los índices de meteorización química son calcio, sodio, magnesio y potasio, ya que sus enlaces químicos son más débiles, en comparación con aluminio y silicio, que tienen enlaces más fuertes y en consecuencia son elementos relativamente inmóviles. En la Tabla 1 se despliegan los porcentajes de los elementos móviles e inmóviles de algunos tipos de roca.

Tabla 1. Resumen de la composición de elementos móviles e inmóviles en % en peso (Yaroshevsky, 2006).

Elementos Roca	calc io	sodi o	magnes io	potasio	alumi nio
Roca básica	6.72	1.94	4.50	0.83	8.76
Roca ultrabásica	0.70	0.57	25.90	0.03	0.45
Roca ácida	1.58	2.77	0.56	3.34	7.70
Meteorito rocoso	1.40	0.70	14.00	0.09	1.30

Dada la amplia variedad de composiciones químicas de las rocas en la superficie terrestre, así como la diversidad de variables ambientales que controlan la movilidad de los elementos, diferentes autores han planteado distintas ecuaciones para calcular índices de meteorización química; algunas de ellas desplegadas en la Tabla 2.

Tabla 2. Fórmulas de los índices de meteorización química con sus respectivas referencias.

Nombre	Fórmula	Referencia
Índice de meteorización de Parker (WIP)	$\left(\frac{Na}{.35} + \frac{Mg}{.9} + \frac{K}{.25} + \frac{Ca}{.7}\right) * 100$	(Parker, 1970)
Índice de Alteración Química (CIA)	$\left(\frac{Al_2O_3}{Al_2O_3 + CaO^* + Na_2O + K_2O}\right) * 100$	(Nesbitt & Young, 1982)
Índice de móviles (I _{MOB})	$\frac{(Ifresco - Imeteorizado)}{Ifresco}$ *I=K ₂ O+Na ₂ O+CaO	(Irfan, 2007)

Parker (1970) plantea el índice de meteorización que lleva su nombre WIP (por las siglas en inglés de *Weathering Index of Parker*), el cual utiliza los elementos de los feldspatos que se movilizan durante la meteorización química y sus coeficientes asociados a la fuerza de enlace en las estructuras minerales. Este índice debe ser utilizado con reserva, ya que puede estimar valores relativamente altos al considerar que todo el calcio en una roca se encuentra únicamente en las estructuras de los silicatos, sin embargo, este elemento puede ser incorporado a los materiales geológicos a través de la precipitación de carbonatos de calcio en cementantes.

Por otra parte, elementos como silicio y aluminio suelen ser considerados inmóviles, no obstante, su inmovilidad no es la misma bajo similares condiciones ambientales superficiales (Formoso, 2006); por ejemplo, el silicio incrementa su movilidad cuando el potencial de hidrógeno del medio es alcalino (pH>7), mientras que el aluminio se mantiene inmóvil bajo condiciones de potencial de hidrógeno dentro de un rango de acidez y basicidad bastante amplio (4<pH<10).

El comportamiento refractario del aluminio en la mayoría de los ambientes geológicos superficiales, así como las consideraciones de la incorporación del elemento

calcio en los procesos de meteorización, derivaron en el planteamiento del índice de alteración química CIA (siglas en inglés de *Chemical Index of Alteration*) de Nesbitt y Young (1982), en el cual se comparan las concentraciones de elementos móviles con el elemento aluminio y se elimina la contribución del elemento calcio de los carbonatos.

Finalmente, el índice de móviles Imob planteado por Irfan (1996, en Irfan 2007) utiliza las concentraciones de elementos móviles en rocas frescas y rocas meteorizadas, brindando resultados óptimos del grado de meteorización química en cuencas con un buen drenaje.

2. Metodología de la fabricación de un juego digital con enfoque educativo

Un juego digital está conformado por un programa que regula el funcionamiento de todo el sistema interno, las reglas, estímulos auditivos, un orden de funcionamiento, los modelos tridimensionales, un objetivo y el diseño visual de la interfaz de usuario (Aktasx & Orcxun, 2016).

La metodología del desarrollo de juegos digitales comerciales, que toma en consideración lo anteriormente descrito, se conforma de las siguientes etapas: la generación del concepto, la pre-producción, la producción, la ejecución de pruebas, el lanzamiento y la post-producción del juego digital (Novak, 2012).

Esta secuencia es generalizada para todo tipo de desarrollo de juegos digitales y en investigaciones se ha desarrollado metodologías dirigidas a la creación de un juego digital con enfoque educativo (Khine, 2010), que será modificada para los objetivos de este trabajo.

En la Tabla 3 se realiza una comparación de la metodología de elaboración de un juego digital comercial propuesta por Novak (2012) y la metodología de desarrollo de juegos digitales con enfoque educativo de Khine (2010).

Tabla 3. Descripción de las fases del desarrollo de juegos digitales comerciales junto al desarrollo con enfoque educativo, con observaciones a los cambios efectuados (Khine, 2010; Novak, 2012).

Fases	Flujo de trabajo		Observaciones
	Actividades a desarrollar para los juegos digitales comerciales	Actividades a desarrollar para los juegos digitales con enfoque educativo	
Concepto	Definición de la idea, mercado, audiencia y plataforma.	Elección del tema en cualquier área, el mercado y audiencia son los alumnos y universidades, se identifica la plataforma idónea.	Definir el concepto es importante para iniciar la pre-producción.
Pre-producción	Definir la historia, crear líneas del tiempo, guion gráfico, un prototipo en papel, diseño de niveles, crear las mecánicas del juego, definir los costos de fabricación, crear un documento del diseño del juego.	Definir si desarrollar una historia. Se diseñan los niveles, las mecánicas de juego, se definen los costos de fabricación y se crea el documento del diseño del juego.	Debido a que se utilizan en el salón de clase, la definición de todos los parámetros no se vuelve obligatoria, para disminuir la longitud del juego.
Producción	Creación de los modelos 3D o 2D, el audio, la interfaz gráfica y la escritura de programación.	Creación de los modelos 3D o 2D, el audio, la interfaz gráfica y la escritura de la programación, tomando en	El uso de las teorías de enseñanza-aprendizaje marca la diferencia en gran medida con el

		cuenta las teorías de enseñanza-aprendizaje.	flujo de trabajo de juegos digitales comerciales.
Ejecución de pruebas	Se efectúan pruebas del funcionamiento para dar con los últimos detalles o fallos mínimos del sistema.	Se realizan pruebas del funcionamiento del juego antes de ser lanzado.	Esta fase no tiene cambios significativos.
Lanzamiento	Se da a conocer el lanzamiento del juego en distintas fuentes de divulgación, con todas las plataformas que se definieron en la fase de concepto.	Se da a conocer el lanzamiento del juego para la implementación en el aula de clase para mejorar la eficacia de la enseñanza y aprendizaje.	Como el enfoque no es enteramente comercial o diversión de un público general, se mantiene el lanzamiento dentro de la audiencia definida en la fase de concepto.
Post-producción	Búsqueda de fallos en el sistema con ayuda de los usuarios y creación de posibles actualizaciones o expansiones del juego.	Búsqueda de fallos en el sistema con ayuda de los usuarios y creación de posibles actualizaciones o expansiones del juego, además de nuevos métodos de unificar el juego en una clase.	De aquí en adelante hay más posibilidades para el juego ya que se pueden dar más ideas de retos, puntuaciones o actividades al llevarlo al salón de clase.

El flujo de trabajo definido para la creación de un juego digital con enfoque educativo es similar al flujo de trabajo de creación de juegos digitales comerciales, pero debido a los objetivos de uso en el salón de clase se delimitaron varias fases de la construcción. Además, cuando el objetivo del juego digital es muy específico se puede reducir el tamaño del proyecto, sobre todo para hacerlo adaptable a los diferentes

tópicos que se requieren enseñar; aunque la mayoría de los juegos comerciales que hay en el mercado necesitan más de 30 horas de juego para su conclusión, para los juegos digitales desarrollados con enfoque educativo se puede reducir la duración del juego a algunos minutos para volverlo casual (Aktax & Orcxun, 2016).

Los juegos casuales son representados como juegos que tienen una jugabilidad simple y fácil de seguir, no cuentan con una continuidad específica para dejar de jugarlo y poder retomarlo en otro momento.

4. Resultado de la aplicación de la metodología propuesta para el tema de índices de meteorización química

Una vez definida la metodología, se llevó a cabo su aplicación para crear un juego digital para uso en el salón de clase, en donde el objetivo es aprender diferentes aspectos del tema de índices de meteorización química.

Concepto. La idea principal del juego digital es visualizar en primera persona los cambios de composición química de las rocas a través del tiempo. La audiencia es de los alumnos de Ciencias de la Tierra que se encuentren cursando la asignatura de Geoquímica General. La plataforma objetivo es el dispositivo móvil, sin embargo, hay posibilidad de adaptación a la PC y tabletas.

Pre-producción. En este caso no se definirá una historia o líneas de tiempo debido a que se requiere que el juego digital sea de corta duración para mantener el interés del estudiante.

Para las mecánicas de juego, se desarrolló la modalidad de juego *Endless Runner*, un tipo de jugabilidad que crea trayectos infinitos con objetivos definidos en el sistema. El principal, y más usado, es el de mantenerse con vida el mayor tiempo posible para acumular los puntajes más altos.

El sistema utiliza parámetros ambientales representativos que controlan la meteorización química, con los que se simulan ambientes naturales como los definidos en la Tabla 4. Los parámetros que se consideraron para cada ambiente son: el relieve, la temperatura promedio y la precipitación anual. Por otra parte, se incluyeron como variables el pH y el tiempo para poder vislumbrar cómo se modifica la movilidad de elementos químicos, la escala temporal en la que un material geológico varía su composición química debido a la meteorización química depende en gran medida de los parámetros ambientales.

Tabla 4. Ambientes naturales y sus respectivos parámetros que lo conforman.

Ambiente natural	Temperatura dominante	Relieve	Precipitación promedio anual
Bosque	Intermedia	Bajo	Media
Desierto	Baja o Alta	Bajo	Baja
Montaña y glaciar	Baja	Alto	Media/Alta
Cueva	Media	Medio	Media
Selva	Alta	Intermedio	Alta

El sistema que controla las partidas dentro del juego contiene un proceso de selección con el que se definen los parámetros de la partida a crear, las elecciones fomentan el sentido de autonomía que describe la teoría de la autodeterminación, el sistema de selección consta de 4 pasos:

El primer paso es la elección de un personaje controlable por el usuario, únicamente con objetivo estético y sin efectos significativos en el transcurso de la partida creada.

El segundo paso es la elección del mapa, con uno de los ambientes mostrados en la Tabla 4, cada opción contiene valores que se utilizan como datos de entrada para modificar los cálculos internos del sistema. Esto modifica la concentración de los elementos químicos con el paso del tiempo en el juego. Durante el desarrollo de una partida cada segundo puede representar 10, 100 o 1,000 años, lo cual define la duración de la partida.

Cabe destacar que cada opción de mapa contiene una lista de partes que conforman la pista de la partida, por ejemplo, un bosque cuenta con árboles, troncos, rocas, suelo, etc., mientras que, un desierto tiene arena, rocas, plantas pequeñas, etc. El objetivo de desarrollar la parte gráfica del mapa es el de otorgar un efecto de inmersión mayor en comparación con solo representar el mismo mapa una y otra vez.

El tercer paso es elegir un objetivo a cumplir como:

- Evitar que la roca seleccionada pierda por completo uno de los elementos químicos para durar en la partida el mayor tiempo posible.
- Movilizar lo más pronto posible uno o todos los elementos químicos de la roca seleccionada.

Los objetivos únicamente implican jugar contra reloj, pero el jugador tiene que utilizar estrategias y conocimiento de la movilidad de los elementos para alcanzar el objetivo seleccionado. Este sistema se puede recorrer una y otra vez para que el usuario pruebe distintas hipótesis de cómo lograr los objetivos, debido a que se ha aclamado que una de las grandes ventajas de los juegos es que fomentan el aprendizaje a través de prueba y error, y que brindan durante el juego información para la retroalimentación (Gee, 2009).

Por último, en el cuarto paso se elige una roca fuente, ya que cada una tiene una composición diferente con la que se inicia el porcentaje de elementos químicos. Estas rocas tendrán valores promedios de los elementos químicos calcio, sodio, potasio y magnesio (Tabla 1).

Durante el paso del tiempo en la partida, el porcentaje de elementos químicos disminuye a una determinada tasa, lo que produce un cambio en los índices de meteorización química con las fórmulas descritas en la Tabla 2, que se calculan en tiempo real y se grafican para que el usuario pueda visualizarlo en el menú de pausa en el juego. Esto tiene el objetivo de que los usuarios descubran que están haciendo bien o que están haciendo mal y poder idear por cuenta propia alguna estrategia para lograr los objetivos elegidos. El sistema de partida termina en victoria si se cumple el objetivo definido en el paso tres o en derrota si el objetivo no se cumple o si el usuario choca tres veces con algún obstáculo de la pista.

El cálculo de los cambios en las concentraciones de elementos químicos es dependiente de las variables del ambiente, que se modifican con objetos potenciadores que provocan eventos en el juego. Los objetos potenciadores aparecen de manera aleatoria de una lista programada en el sistema, cada objeto provoca un evento diferente. Los eventos modifican algún parámetro del ambiente, cuando el usuario recoge un objeto en forma de sol produce un incremento en la temperatura del ambiente o viceversa si el objeto recogido tiene forma de copo de nieve. Otro evento modifica el valor de precipitación, el cual sube cuando se recoge un objeto con forma de gota y al mismo tiempo el pH se vuelve más alcalino.

Durante una partida, el jugador debe formular una estrategia para lograr sus objetivos, cada objeto potenciador puede ser beneficioso o perjudicial según los objetivos de la partida, así que, el usuario puede decidir si agarra alguno de estos objetos potenciadores o no.

Producción. El juego digital se creó en el software libre de *Unity*, con el lenguaje de programación de *C#*, con el apoyo de una plantilla gratuita del formato de *Endless Runner*.

Las distintas fases del juego digital se presentan más adelante en una serie de imágenes. En la Figura 1 se ilustran los primeros pasos definidos en la preproducción, la selección del personaje, la selección de mapa y el menú de selección de modos.

La selección del mapa contiene una sección de visualización de parámetros, la temperatura, precipitación y relieve iniciales están representados como barras que muestran de manera gráfica su valor. El menú de modos tiene una sección de descripción para explicar cómo funciona el modo seleccionado y algunos consejos para lograr el objetivo.



Fig. 1. Ilustraciones de las distintas interfaces de usuario de las fases de selección que tiene el sistema, a) elección del personaje, b) menú de selección de mapa y c) menú de elección de modos.

La siguiente fase es de la selección de rocas con una sección de visualización gráfica de los valores de concentración de elementos químicos de cada roca. En este menú se muestra en una escala el valor de cada elemento y muestra de manera gráfica el nivel de dificultad de la partida, dependiendo de los parámetros seleccionados en los diferentes menús del sistema, se ejemplifica en la Figura 2a.

La interfaz gráfica del juego en la partida muestra diferentes contadores de puntaje, años transcurridos, vidas en la partida, puntajes gráficos de elementos químicos y un contador gráfico de la temperatura de la partida en curso, se ejemplifica en la Figura 2b. Los contadores en la interfaz ayudan a disminuir la carga de la memoria de trabajo, otorgando la información relevante en forma de imágenes sencillas y ayuda al usuario a concentrarse en el objetivo de la partida.



Fig. 2. Continuación de las ilustraciones de las fases del sistema, a) menú de selección de roca y medición de dificultad de la partida, b) interfaz gráfica con contadores de puntajes en tiempo real de la partida y c) menú de pausa.

En el menú de pausa de partida se dan las opciones de salir de la partida, continuarla

y de abrir el menú de gráficas. La interfaz de las gráficas contiene diferentes botones que cambian el menú para visualizar las diferentes variables de la partida como se muestra en la Figura 3, en donde se observan los datos presentados en dos formas: gráfica de líneas y gráfica de barras.

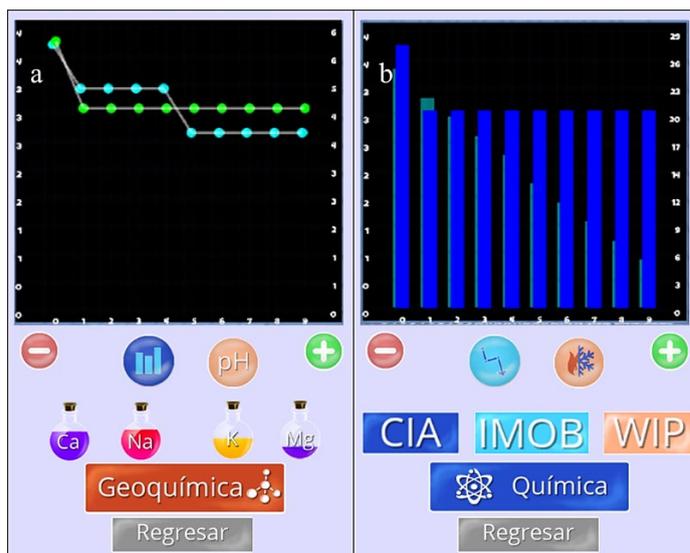


Figura 3. Interfaz gráfica para la visualización de las variables de la partida conformado por a) menú de selección de los datos de elementos químicos en formato de gráfica de líneas, b) menú para la selección de las variables de índices de meteorización para graficar en formato de grafica de barras, en ambos gráficos en el eje de las ordenadas se despliega la leyenda del parámetro seleccionado y en el eje de las abscisas el tiempo transcurrido en la partida.

Para la creación de mapas, el sistema hace una integración de elementos visuales en una línea recta, se aplica una ligera curvatura para esconder el final del trayecto y se logre la ilusión de que no tiene un fin. Algunos ejemplos de estos mapas están en la Figura 4, cada mapa tiene una diferencia significativa en los elementos visuales que lo conforman, dando al usuario un sentido de mayor inmersión.

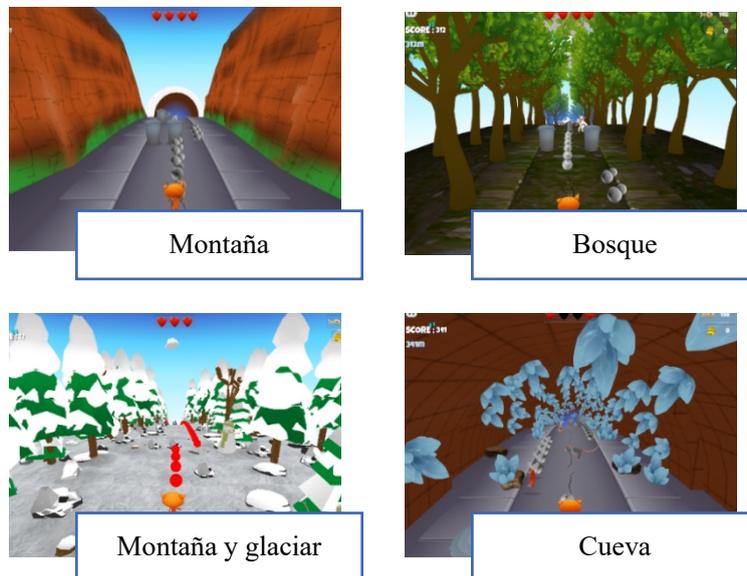


Fig. 4. Ilustraciones de los diferentes ambientes que el prototipo puede simular en el juego con sus respectivos nombres, cada uno representando diferentes variables.

El usuario debe pensar qué variables modificar en el ambiente para lograr los objetivos de la partida, usando todos esos cambios a su favor. Un ejemplo es subir la temperatura y las lluvias en un bosque para hacer que los elementos se movieran más rápido o viceversa si se quiere resistir por más tiempo en la partida.

5. Conclusión

Con base en la evidencia, la educación tiene muchas posibilidades para volverse más eficaz, la implementación de una metodología para la creación de juegos digitales educativos es una oportunidad para innovar la enseñanza universitaria y para mejorar la transferencia del conocimiento.

Las teorías de enseñanza-aprendizaje apoyan en gran medida la utilización de juegos digitales en el salón de clase, es por ello que las universidades deben implementar una metodología que tome en cuenta todos los avances en esta área.

En este trabajo se utilizó la metodología para examinar el proceso de desarrollo de juegos digitales con enfoque educativo, con el área objetivo de Ciencias de la Tierra, a su vez la implementación de la metodología propuesta tiene la oportunidad de ser aplicada a otros temas fuera de las Ciencias de la Tierra para transformar el proceso de enseñanza-aprendizaje. Es importante recalcar que la construcción de los juegos digitales requiere de habilidades específicas para mejorar la calidad de los resultados y disminuir el tiempo de construcción.

6. Referencias

1. Aktasx, A. Z., & Orcxun, E. (2016). A survey of computer game development. *Journal of Defense Modeling and Simulation*, 13(2), 239–251. <https://doi.org/10.1177/1548512914554405>
2. Ashinoff, B. (2014). The potential of video games as a pedagogical tool. *Frontiers in Psychology*, 5(SEP), 1–5. <https://doi.org/10.3389/fpsyg.2014.01109>
3. Becker, K. (2017). *Choosing and Using Digital Games in the Classroom*. <https://doi.org/10.1007/978-3-319-12223-6>
4. Blunt, R. (2007). Does Game-Based Learning Work? Results from Three Recent Studies. *Advanced Distributed Learning*, 945–954. <https://doi.org/10.1.1.502.4585>
5. Bruner, J. (1961). The Act of Discovery. *Harvard Educational Review*. [https://doi.org/10.1016/S0040-4039\(02\)00317-9](https://doi.org/10.1016/S0040-4039(02)00317-9)
6. Crocco, F., Offenholley, K., & Hernandez, C. (2016). A Proof-of-Concept Study of Game-Based Learning in Higher Education. *Simulation and Gaming*, 47(4), 403–422. <https://doi.org/10.1177/1046878116632484>
7. David Merrill, M. (2009). First principles of instruction. In *Instructional-Design Theories and Models*. <https://doi.org/10.4324/9780203872130>
8. El Mawas, N., & Muntean, C. H. (2018). SUPPORTING LIFELONG LEARNING THROUGH DEVELOPMENT OF 21ST CENTURY SKILLS. In *EDULEARN18 Proceedings*. <https://doi.org/10.21125/edulearn.2018.1723>
9. Formoso, M. L. L. (2006). Some topics on geochemistry of weathering: A review. *Anais Da Academia Brasileira de Ciencias*. <https://doi.org/10.1590/S0001-37652006000400014>
10. Gagne, R. (2008). Conditions of Learning (R . Gagne). *Learning*.
11. Gee, J. P. (2009). Deep learning properties of good digital games: How far can they go? In *Serious Games: Mechanisms and Effects*. <https://doi.org/10.4324/9780203891650>
12. Illeris, K. (2018). An overview of the history of learning theory. *European Journal of Education*. <https://doi.org/10.1111/ejed.12265>
13. Irfan, T. Y. (2007). Characterization of weathered volcanic rocks in Hong Kong. *Quarterly Journal of Engineering Geology and Hydrogeology*. <https://doi.org/10.1144/gsl.qjeg.1999.032.p4.03>
14. Khine, M. S. (2010). *Learning to play: Exploring the future of education with video games*. Springer.
15. Li, M. C., & Tsai, C. C. (2013). Game-Based Learning in Science Education: A Review of Relevant Research. *Journal of Science Education and Technology*, 22(6), 877–898. <https://doi.org/10.1007/s10956-013-9436-x>
16. Li, Y. S., Zhao, T. S., & Yang, W. W. (2010). Measurements of water uptake and transport properties in anion-exchange membranes. *International Journal of Hydrogen Energy*. <https://doi.org/10.1016/j.ijhydene.2010.03.026>
17. Misra, K. C. (2012). *Introduction to geochemistry: principles and applications*. Wiley-Blackwell.
18. Nesbitt, H. W., & Young, G. M. (1982). Early proterozoic climates and plate motions inferred from major element chemistry of lutites. *Nature*, 299(5885), 715–717. <https://doi.org/10.1038/299715a0>
19. Novak, J. (2012). *Game Development Essentials Third Edition*. _Imp03.

20. Oanh, D. T. K. (2018). Organizing Experiential Learning Activities for Sustainable Development of Core Competences of Technical Students at Ho Chi Minh City University of Technology and Education. In *Proceedings 2018 4th International Conference on Green Technology and Sustainable Development, GTSD 2018*. <https://doi.org/10.1109/GTSD.2018.8595516>
21. Parker, A. (1970). An Index of Weathering for Silicate Rocks. *Geological Magazine*, 107(6), 501–504. <https://doi.org/10.1017/S0016756800058581>
22. Pivac, M., Dziabenko, O., & Schinnerl, I. (2003). Aspects of Game Based Learning.
23. Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of Game-Based Learning. *Educational Psychologist*, 50(4), 258–283. <https://doi.org/10.1080/00461520.2015.1122533>
24. Qian, M., & Clark, K. R. (2016). Game-based Learning and 21st century skills: A review of recent research. *Computers in Human Behavior*, 63, 50–58. <https://doi.org/10.1016/j.chb.2016.05.023>
25. Reigeluth, C. M. (1999). What Is Instructional-Design Theory and How Is It Changing? In *Instructional-Design Theories and Models Volume II: A New Paradigm of Instructional Theory*. <https://doi.org/10.1007/s13398-014-0173-7.2>
26. Roy, P. D., Chávez-Lara, C. M., Beramendi-Orosco, L. E., Sánchez-Zavala, J. L., Muthu-Sankar, G., Lozano-Santacruz, R., ... López-Balbiaux, N. (2015). Paleohydrology of the Santiaguillo Basin (Mexico) since late last glacial and climate variation in southern part of western subtropical North America. *Quaternary Research (United States)*. <https://doi.org/10.1016/j.yqres.2015.10.002>
27. Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*. <https://doi.org/10.1037/0003-066X.55.1.68>
28. Seel, N. M., Lehmann, T., Blumschein, P., Podolskiy, O. A., Seel, N. M., Lehmann, T., ... Podolskiy, O. A. (2017). What is Instructional Design? In *Instructional Design for Learning*. https://doi.org/10.1007/978-94-6300-941-6_1
29. Shick, J. M., Hoffmann, R. J., Lamb, A. N., & Hole, W. (2015). The Relative Solubilities of the Chemical Constituents of Rocks, 19(3), 699–713. <https://doi.org/10.1002/jae.1083>
30. Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*. [https://doi.org/10.1016/0364-0213\(88\)90023-7](https://doi.org/10.1016/0364-0213(88)90023-7)
31. Tarbuck, E., & Lutgens, F. K. (2005). *Ciencias de la tierra*. Pearson (8th ed., Vol. 12). [https://doi.org/10.1016/0196-0709\(91\)90073-O](https://doi.org/10.1016/0196-0709(91)90073-O)
32. Yaroshevsky, A. A. (2006). Abundances of chemical elements in the Earth's crust. *Geochemistry International*, 44(1), 48–55. <https://doi.org/10.1134/s001670290601006x>

Ambiente de coworking y co-creación en red para fortalecer el proceso de aprendizaje e investigación en la formación académica-científica. Caso UTN

Alexander Guevara Vega^{a,b}, Miguel Naranjo Toro^a, Cathy Guevara Vega^{a,b}, Andrea Basantes Andrade^a,
Sania Ortega Andrade^{a,c}

^a Universidad Técnica del Norte, Grupo de Investigación de Ciencias en RED eCIER, Antiguo Hospital
San Vicente de Paúl, Calle Juan Montalvo 8-46, Ibarra, Ecuador
alexguevara@utn.edu.ec, menaranjo@utn.edu.ec, avbasantes@utn.edu.ec, cguevara@utn.edu.ec,
smortega@utn.edu.ec

^b Universidad Técnica del Norte, Facultad de Ingeniería en Ciencias Aplicadas FICA, Av. 17 de Julio
5-21, Ibarra, Ecuador
alexguevara@utn.edu.ec, cguevara@utn.edu.ec

^c Universidad Técnica del Norte, Facultad de Ingeniería en Ciencias Ambientales y Agropecuarias
FICAYA, Av. 17 de Julio 5-21, Ibarra, Ecuador
smortega@utn.edu.ec

Resumen. El coworking en las Instituciones de Educación Superior (IES) ha generado nuevos espacios de participación entre docentes y estudiantes encaminado a un nuevo cambio de paradigma en los procesos de aprendizaje e investigación, esto conlleva a que las Tecnologías de Información y Comunicación (TIC) sirvan de apoyo en el fortalecimiento de la nueva era colaborativa en las IES. El objetivo es crear una cultura de aprendizaje colaborativo e investigación formativa en el aula, dinamizando un mayor nivel de transferencia de ideas, las relaciones sociales basadas en la colaboración y la co-creación en red para trabajar en equipo, generando una mejor interacción entre docentes y estudiantes a través de herramientas de comunicación, colaboración y productividad robustas en la nube, que permita minimizar recursos económicos, elevando la productividad de las tareas en menor tiempo, agilizando los procesos académicos y de investigación. Para ello se desarrolló un ambiente de hibridación y co-creación mediante la creación del grupo de investigación eCIER de la Universidad Técnica del Norte (UTN) y la implementación de un ambiente unificado a medida de comunicación, colaboración y productividad en la UTN, su despliegue curricular de contenidos del modelo de coworking se realizó utilizando Design Thinking & Agile y para la implementación tecnológica FastTrack como metodologías. El nuevo ecosistema de aprendizaje basado en investigación y el modelo de co-creación en red en la UTN permitió mejorar el proceso de investigación dentro y fuera del aula, la comunicación, colaboración y productividad entre los involucrados fue más eficiente en un 80%, dotando de un ambiente de aprendizaje e investigación ágil y dinámico, disminuyendo notablemente la brecha digital e incrementado la generación de nuevo conocimiento y el uso de recursos digitales accesibles manteniendo una fluidez tecnológica entre la UTN, las redes de investigación y proyectos de cooperación internacional.

Palabras Clave: Coworking, co-creación, colaboración, comunicación, productividad, investigación, eCIER, innovación, aprendizaje, design thinking, agile.

Eje temático: Tecnologías en la innovación educativa e investigación.

1. Introducción

El proceso de investigación formativa e innovación educativa está presente en los nuevos modelos de ambientes académicos colaborativos en las Instituciones de Educación Superior (IES) con aulas modernas, interactivas e inmersivas de coworking y talento humano capacitado, innovador, creativo y tecnológico. La universidad en el Ecuador se encuentra en una transición digital, hiperconectada y colaborativa de forma casi natural, el trabajo en red y la co-creación entre los miembros de la comunidad, afianza nodos y vínculos de cooperación que se crean con ellos a través de la experiencia, a cada actor le gusta ser protagonista por ello su experiencia de usuario (UX) es mucho más enriquecedora y se aprovecha mejor el potencial innovador que puede ser transmitido al aula, generando ideas de cambio en los modelos de aprendizaje tradicionales y enfocándose en un modelo de conectivismo creativo y sobre todo accesible y abierto

El acciona de las nuevas exigencias de la industria más competitivas y de mejora continua ha obligado a las universidades generar cambio en sus profesionales, dotándoles de nuevas habilidades y competencias, así como la incorporación del uso de las TIC en la educación superior permite integrar un conjunto de acciones, estrategias y herramientas para hacer visible el cambio orientado y encontrar la excelencia académica en las IES. En el Ecuador (fuente INEC 2017) el 10,5% de las personas de 15 a 49 años en el Ecuador son analfabetas digitales, 10,9 puntos menos que en el año 2012 [3].

Los estudiantes de la actualidad necesitan que los procesos de aprendizaje e investigación se integren y generen valor y cambio, la cultura de trabajo en red toma en cuenta a personas de una generación digital, el aprendizaje debe fomentar la colaboración, resolución de problemas, pensamiento computacional, uso de las TIC y demás habilidades del siglo 21 [2], [3].

La Universidad Técnica del Norte (UTN) se propuso adoptar un nuevo modelo de aprendizaje basado en investigación, y trabajo en red enlazado a un esquema tecnológico integrado de comunicación y colaboración avanzada e interactiva, que le permita minimizar problemas como: limitados espacio físicos para crear, debatir, colaborar, compartir conocimiento, la pérdida de información y conversaciones gracias a la mensajería persistente, dispersión de contenidos en diferentes herramientas y una inadecuada retroalimentación entre docentes y estudiantes; todo esto llevó a considerar un cambio de ecosistema en una comunidad emergente del cambio, en busca de la innovación, pasar de ambientes físicos tradicionales a modernos espacios digitales inmersivos e interactivos con múltiples recursos, logrando mejorar la UX.

El presente trabajo da a conocer, una propuesta de hibridación y co-creación mediante la creación del grupo de investigación eCIER y la implementación de un ambiente tecnológico de comunicación y entornos de colaboración en red utilizando la plataforma Cloud con servicio SaaS (Software as at Services) Microsoft Teams disponible como solución en el campus [4], permitiendo a docentes y estudiantes de la UTN definir nuevas experiencias de aprendizaje e investigación dentro y fuera del aula [5].

2. Antecedentes

2.1 Necesidad de crear nuevos espacios de coworking y co-creación de nuevas ideas a través de plataformas tecnológicas para la UTN

El Ecuador en los últimos años le ha apostado a la generación de investigación y conocimiento gracias a políticas públicas y a un alto porcentaje de inversión en investigación y desarrollo por parte del gobierno central, a mediados de 2019 los expertos comentan que Ecuador tiene un problema en cuanto a la inversión: solo se dedica a investigación y desarrollo el 0,4 % del PIB, por el 2,4 % de media que dedican los países de la Organización para la Cooperación y el Desarrollo Económicos (OCDE) [6].

La estrategia radica en que las universidades actualmente han triplicado su inversión en ciencia y tecnología y superan el porcentaje mínimo que se contempla en la Ley Orgánica de Educación Superior (LOES) que es el 6%. Los nuevos estándares de calidad en la educación a los que se rigen las IES han permitido generar acciones y planes de mejoramiento continuo en busca de una educación de calidad [6].

Las nuevas tendencias se enfocan a una innovación más participativa, colaborativa y sobre todo multidisciplinaria en beneficio de un solo objetivo, "el de fortalecer la educación", a esto se suma iniciativas de capital semilla por parte de instituciones de investigación como SENESCYT, CEDIA, FRIDA, UNESCO, VLIR-OUS, ONU como ejemplo y la empresa privada que busca en los centros de educación superior la solución a los problemas mundiales y en muchos de los casos locales.

El impacto de la desmaterialización es uno de los aspectos principales para la reducción de las emisiones de GtCO₂e, en este caso se enfoca en la desmaterialización en el aula que puede aplicarse a diversos aspectos de la vida cotidiana o trabajo habitual por sus siglas en inglés Business As Usual (BAU); con el objetivo de reducir el número de objetos materiales que necesitan producirse para una clase, una reunión, una capacitación o generación de una nueva idea. Las reuniones de co-creación en red, trabajo colaborativo, los documentos electrónicos, libros online, audio y video digital, en sustitución de reuniones presenciales, duplicidad de información, el uso agresivo del papel y los DVD, reducirán las emisiones asociadas con su fabricación y distribución. La Fig. 1 muestra el impacto de la desmaterialización – GtCO₂e estudio realizado por The Climate Group Smart 2020 report elaborado en 2008 por Telefónica [7].



Fig. 1. El impacto de la desmaterialización – GtCO₂e, según el reporte The Climate Group Smart 2020.

2.2 Fortalecer el proceso de aprendizaje basado en la investigación y la productividad en el aula

Diseñar un ciclo de investigación de aula formativa desde el proceso de investigación institucional y científico a un modelo de investigación en el aula, para ello se tiene como base la Taxonomía Digital de Bloom (1956) y Revisores Anderson/Krath Wohl (2001) y Churches (2009), la investigación llevada al aula permite al estudiante ser un actor dentro del proceso haciendo participe de sus ideas y permitiéndole que construya su propio conocimiento con tutelas presenciales y actividades online, generando en el docente la competencia de mentor de cambio en cada estudiante [8].

Cuando se habla del término productividad según la Real Academia Española (RAE) se define como: “*La relación entre lo producido y los medios empleados*”, la productividad suele estar asociada a la eficiencia y al tiempo, cuanto menos tiempo se invierta en lograr el resultado anhelado, mayor será el carácter productivo del sistema [9].

Ambos contextos aportan elementos relacionados a la investigación y aprendizaje productivo dentro y fuera del aula, esto se puede extrapolar al estudio de caso de la UTN en el cual se debe saber: la relación entre la generación de conocimiento producido y los medios empleados, que es útil y provechoso, la capacidad o grado de producción y que resultados de valor se obtuvieron, de esta manera se puede definir que el mejoramiento del aprendizaje basado en investigación y la productividad en el aula es la manera en cómo incide el factor humano y tecnológico para gestionar con eficiencia y eficacia todos los recursos disponibles, con el fin de obtener resultados altamente positivos facilitando no solo la gestión de recursos si no en el fortalecimiento del conocimiento del estudiante de la UTN [10].

La teoría del conectivismo en las competencias del siglo 21 en las IES se acerca a disponer de una variedad de herramientas robustas pero que en la práctica no son aprovechadas al máximo, muchas veces esa explosión se deriva del desconocimiento

que tiene su implementación en espacios de co-creación y más bien se enfoca en cómo utilizar de manera tradicional en el proceso de aprendizaje e investigación, lo que provoca una baja productividad en la comunidad UTN. Por lo que, experimentar nuevos ambientes como desafío y romper con viejos métodos de enseñanza es el objetivo de las IES, permitiendo motivar al auto aprendizaje de una forma colaborativa construyendo su propio conocimiento e incrementando una autonomía académica con responsabilidad y que genere valor [10].

2.3 Análisis tecnológico

El término de Cloud (nube) se refiere a la forma de ver a una red de computadoras interconectadas que brindan un aprovisionamiento de servicios de software y datos, el sector educativo, entendido en un sentido amplio, es cada vez más dependiente de las actividades a distancia (educación o formación a distancia), por lo que la utilización de las posibilidades y servicios que ofrece la nube se ha convertido en algo necesario y estratégico, por lo que el uso y aplicación del Cloud Computing en las universidades afecta tanto a los propios procesos académicos como a los de investigación.

El Cloud Computing en las IES y en el caso en particular en la UTN forma parte de uno de sus ejes estratégicos en la innovación educativa, brindando la posibilidad de co-crear en red y generar espacios de coworking entre nodos y redes de investigación a gran escala, la alta disponibilidad de los servicios y contenidos permite generar agilidad, confianza y sobre todo motivación entre la comunidad simplificando las tareas de aprendizaje e investigación, adquiriendo competencias modernas y digitales, el resultado de todos estos elementos es que los sistemas y plataformas educativas están cada vez migrando a la nube [11].

La colaboración y comunicación entre participantes, se considera como punto de partida en las tecnologías de Cloud Computing, las plataformas y entornos desarrollados disponen de herramientas de comunicación que bien usadas, facilitan y estrechan la relación entre todos los agentes que participan en los procesos de investigación y académicos [11].

3. Desarrollo e implementación de un ambiente de coworking y co-creación en red para fortalecer el proceso de aprendizaje e investigación en la formación académica-científica. Caso de la Universidad Técnica del Norte

3.1 Necesidad de crear nuevos espacios de coworking y co-creación de nuevas ideas a través de plataformas tecnológicas para la UTN

Una de las estrategias para fortalecer el proceso de investigación y el aprendizaje basado en investigación en el aula fue el de crear el grupo de investigación de ciencias en red eCIER UTN que tiene por objetivo promover, desarrollar y estimular proyectos

que permitan fortalecer los procesos de investigación aplicando buenas prácticas en el uso de las TIC bajo el paradigma de construir una Sociedad del Conocimiento y la Información como eje transversal como aporte para la comunidad científica UTN.

El grupo eCIER - UTN busca sinergias de trabajo colaborativo, coworking y co-creación con redes y grupos de investigación internos a la UTN y externos a nivel nacional e internacional, actualmente dispone de 36 miembros activos en la academia de diferentes áreas del conocimiento, 16 miembros asociados externos y de 10 redes y grupos académicos y de investigación asociados, desde su creación en noviembre del 2017 con base y asesoría de Red CLARA y la comunidad RED eCIENCIA logra constituirse como el grupo con más proyectos con fondos concursables, con más participaciones de coworking y co-creación en red entre la UTN con 3 continentes (América, Europa y África), la movilidad estudiantil mediante sus proyectos de investigación y el aprendizaje basado en la investigación son los instrumentos que han promulgado una cultura de innovación en la UTN gracias a la cooperación internacional.

La gestión y seguimiento de los proyectos de aula con el aprendizaje basado en la investigación son encubados gracias a la aplicación de metodologías de diseño instruccional basadas en design thinking y la ejecución ágil con SCRUM, el apoyo de tecnologías son de vital importancia por ello el uso de Microsoft Teams han permitido integrar el modelo de investigación formativa y co-creación en red, esta solución dispone de canales de comunicación persistentes y de soporte de múltiples conexiones a gran escala, soportadas por videoconferencias en HD, colaboración, mensajería, colaboración de documentos con edición simultáneas así como la gestión de actividades mediante un tracking sostenido por componentes de terceros como plugin adicionales para su funcionamiento, esto hace de una solución robusta todo en uno.

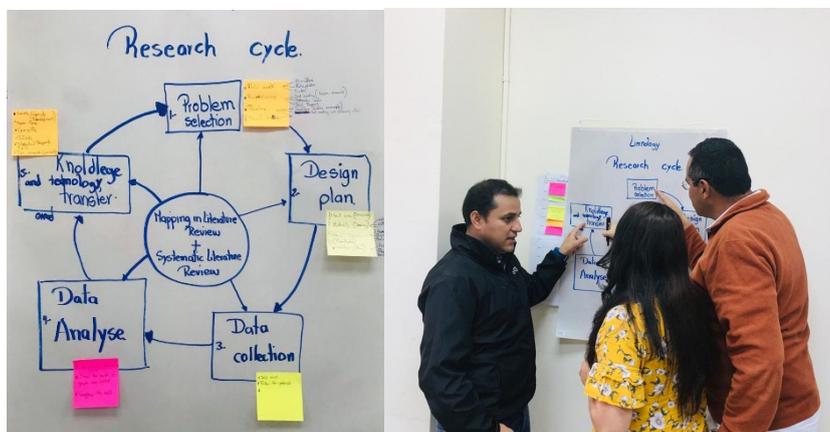


Fig. 2. Aplicación de coworking y co-creación – Grupo eCIER UTN, capacitación de aprendizaje basado en investigación, proyecto Network VLIR-OUS Ecuador.



Fig. 3. Participación de clase compartida entre el grupo eCIER UTN y con 3 continentes (América, Europa y África), uso de Teams como plataforma de colaboración.

3.2 Implementación del ambiente de comunicaciones avanzadas UTN mediante la metodología FastTrack

La solución de Microsoft Teams de Office 365 permite que la organización realice el trabajo de forma más conectada, más estructurada, más natural, más inteligente, más móvil y con más colaboración. Para ello se estableció un estudio previo aplicando el concepto de evolución en la productividad, mismo que pasa de un modelo individual a un modelo empresarial On Premise y finalmente a un modelo Cloud. En este tipo de ambiente es necesario contar con un plan de despliegue y adopción ágil con una visión de cultura organización de alto nivel que genere valor a la institución.

Para la implementación del ambiente de comunicaciones avanzadas se aplicó la Guía de Adopción de trayectoria rápida orientada a productos y servicios tecnológicos, basada en la metodología de trabajo nativa de Microsoft "FastTrack", que tiene las siguientes fases: Visión, Incorporación y Generar Valor [12].

En la fase de Visión, se seleccionó a los usuarios que tienen carga horaria de trabajo autónomo se seleccionó a 60 personas entre ellos 30 docentes e investigadores de eCIER, 30 estudiantes de pregrado con competencias básicas de TIC, con ellos se estableció los escenarios necesarios para captar la aplicabilidad de la plataforma en el menor tiempo posible creando un plan de adopción que tenga éxito [12].

En la fase de Incorporación, se consideró las variables de gestión del cambio y aplicabilidad de la plataforma al trabajo autónomo de cada asignatura con los conocimientos necesarios de los usuarios avanzados y con el 99% de aceptación en el uso y manejo de Microsoft Teams [12].

La fase de Generar Valor, permitió generar un valor y mejora continua de la implementación con la participación activa del usuario que permitió impulsar el 99,9% de adopción, evidenciando el cambio y reconociendo los errores. La figura 5, muestra las fases de la guía de adopción de la metodología FastTrack [12].

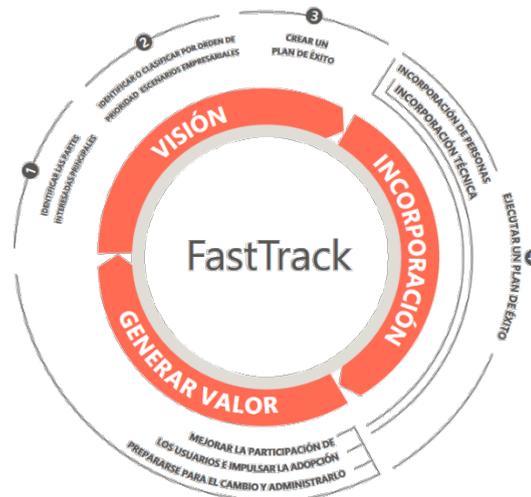


Fig. 4. Guía de adopción de trayectoria rápida (FastTrack).

Fuente: <http://fasttrack.microsoft.com>

Para dar cumplimiento a las fases en el proceso de adopción, se planificó cursos de capacitación guiadas de tipo workshop y talleres claves de tipo keynote, al personal docente se realizó mini workshop denominados: "La nueva era de las comunicaciones avanzadas en el aula", con una duración de 12 horas en modalidad presencial y autónomas, controladas y asistidas mediante la plataforma de Microsoft Teams; que facilitó la inducción, visión global del entorno, entrenamiento en las herramientas de valor, seguimiento, evaluación y lecciones aprendidas. Para la adopción de los estudiantes se desarrolló un keynote denominado: "Innovación digital en el aula: @utn.edu.ec como marca" con dos sesiones diarias por grupo, de 4 horas de capacitación durante 4 días de fin de semana en modalidad virtual mediante la plataforma de Microsoft Teams.

El seguimiento y validación del proceso de adopción permitió crear indicadores de valor para la ejecución del trabajo autónomo mismo que se encuentra integrado al plan de clase de cada asignatura y a la vez permitió desarrollar habilidades y competencias digitales en los docentes y estudiantes de la UTN, con la adopción progresiva ayudarán a masificar la aplicabilidad de Microsoft Teams en todo el campus; Las redes académicas y de investigación participantes formaron parte de la adopción y apoyaron a esta iniciativa de Transformación Digital.

4. Resultados

4.1 Métricas cualitativas y cuantitativas de las estrategias de comunicaciones avanzadas dentro y fuera del aula

Desde que se realizó la implementación de Office 365 en la UTN en el año 2014, en esta solución se han activado 28.804 cuentas de usuarios, adicionalmente cuentan con el servicio de Microsoft Teams, es decir que el 100% de la comunidad universitaria está dentro del ambiente tecnológico, aquí se incluyen a los 654 docentes, 10.784 alumnos, 16.881 exalumnos, 413 administrativos y 72 directivos.

La UTN ha reducido considerablemente el gasto operativo, de equipamiento y la inversión en tecnología para disponer de servicios de comunicación y colaboración a nivel académica, consecuente con esto ha permitido fortalecer la cobertura del servicio de internet a 1000 Mb, fortalecer los laboratorios de TIC, así como las terminales de acceso común en zonas de consulta con accesorios para videoconferencias como cámaras y headphones, además de incorporar un plan de capacitación de TI al personal que brinda el soporte y mantenimiento de la plataforma.

Las métricas cualitativas tienen un enfoque de conocer el valor que ha marcado en el fortalecimiento de la productividad en sus procesos académicos dentro y fuera del aula, existe un alto grado de aceptación de Microsoft Teams en el aula ya que ayuda en gran parte a minimizar el tiempo que co-creación en red, así como el proceso de tutoría de investigación en los programas de maestría como de pregrado, lo que a permito ser más eficientes y productivos tanto docentes y estudiantes, en el caso del aprendizaje basado en la investigación ha permitido adquirir compromisos y actitudes de valor tanto en sus docentes ya que son más productivos en la planificación de actividades propias de la investigación como el de desarrollar habilidades de coworking llevando a un enfoque motivacional ya que el estudiante no necesita estar presente bajo presión del mismo ambiente del aula sino más bien desde cualquier lugar de preferencia y mediante el dispositivo que más le agrada puede desarrollar con total tranquilidad este proceso de aprendizaje, ayudando a fortalecer el proceso de investigación en la UTN.

Las métricas cuantitativas están determinadas en: la reducción del 60% en el uso de herramientas de comunicación asíncronas como el correo electrónico, la reducción del 80% en el número de reuniones no planificadas y la reducción del 90% en las solicitudes de soporte a TI en cada estación de trabajo. Además, el 85% de mejora en el ambiente académico y espíritu de equipo fomentando una cultura colaborativa en la comunidad universitaria en especial en los grupos de investigación multidisciplinarios, el 90% de mejora en la gestión de tareas y actividades compartidas de forma colectiva gracias al sistema de comunicaciones avanzadas permitiendo minimizar en un 50% el impacto ambiental que provoca el gasto de movilización, a través del uso de una plataforma de videoconferencias institucional.

Para iniciar con el plan de adopción de la plataforma se realizó una evaluación de habilidades y destrezas durante los mini workshops y keynote mediante un test anónimo de competencias digitales, en el cual se obtuvo como resultado, el 99% de los resultados se situaban entre 7 y 9.5, considerado como rango aceptable el desarrollo del conocimiento y destrezas adquiridas en el periodo de adopción de la plataforma de

comunicaciones avanzadas. Esto demostró que los usuarios estaban preparados para la transformación digital y convertir sus clases en aulas móviles, modernas e interactivas, un claro ejemplo de la aplicación del trabajo autónomo y el aprendizaje inmersivo.

Actualmente se ha implementado un plan de capacitación continua para la comunidad universitaria, esto garantiza que el proceso de adopción tecnológica sea más activo y eficiente en la educación, alcanzando sustancialmente una cultura de transformación digital en el aula. Los beneficios de Microsoft Teams en la UTN ha permitido:

- Adoptar a Microsoft Teams como una plataforma de comunicación avanzada de conferencias web, mensajería instantánea, audio y video llamadas y herramienta de colaboración sincrónica, con un nivel de utilidad promedio de 2.800 minutos de audio y video, con periodicidad mensual desde enero de 2017 a marzo de 2019.
- Los investigadores de eCIER fueron los pioneros en adoptar esta herramienta, durante el semestre académico a través de proyecto de co-creación en red, logrando enlazar el sistema de comunicación entre investigadores.

A continuación, se muestran los resultados obtenidos en la implementación de la plataforma. Fig. 5, 6 y 7.

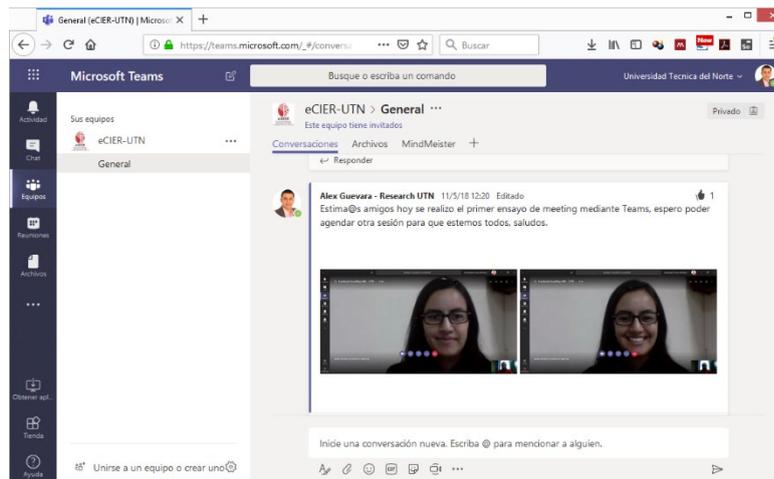


Fig. 5. Despliegue de Teams como plataforma de co-creación, preparación de capacitaciones en el uso de TIC y gestión de proyectos con miembros del grupo de investigación eCIER UTN.

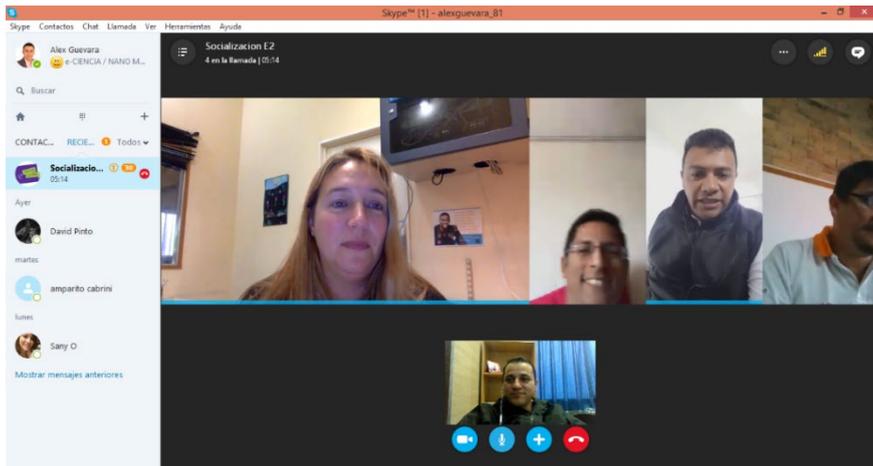


Fig. 6. Participación de coworking entre el grupo de investigación eCIER UTN , profesores de las Red Iberoamericana de Educadores Innovadores de Guatemala, Perú y Argentina.

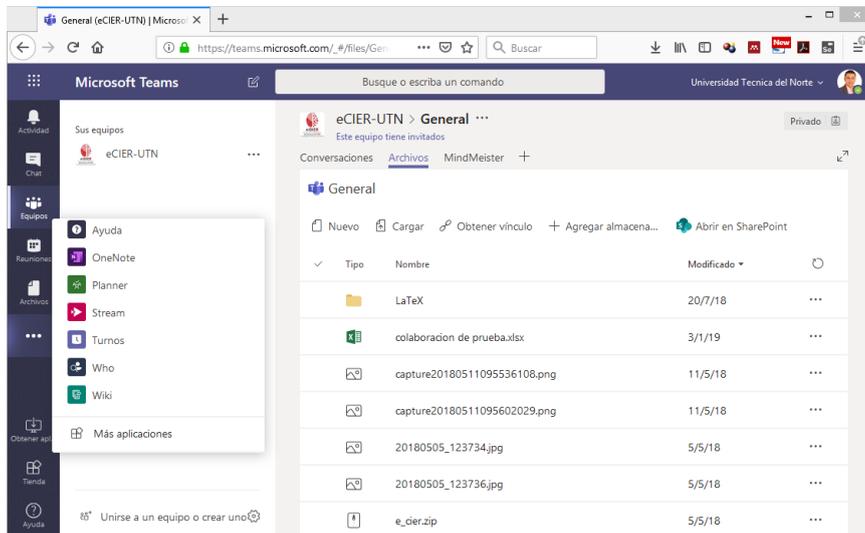


Fig. 7. Integración de componentes extras a la plataforma de Teams para gestionar proyectos de investigación formativa en el aula mediante el uso de recursos hipermediales.

Agradecimientos

Este trabajo ha sido desarrollado gracias a la colaboración directa del Grupo de Investigación de Ciencias en RED eCIER UTN, el equipo técnico de la Dirección de Desarrollo Tecnológico e Informático UTN, la carrera de Ingeniería en Sistemas Computacionales CISIC UTN, y en especial agradecimiento al Dr. Miguel Naranjo Toro por su participación y apoyo en la implementación del trabajo.

Referencias

- 1** La transformación digital del sector educación, http://www.fundacionorange.es/wp-content/uploads/2016/11/eE_La_transformacion_digital_del_sector_educacion-1.pdf
- 2** 4º Congreso Internacional de Innovación Educativa, <http://ciie.itesm.mx/en/platicas-tecnologicas/>
- 3** Instituto Nacional de Estadística y Censos INEC, <http://www.ecuadorencifras.gob.ec/tecnologias-de-la-informacion-y-comunicacion-tic/>
- 4** M. F. Sakr, "Introduction to Cloud Computing Lecture Motivation ...," pp. 15–25; Creación:2010; Recuperado:9 junio 2016, 2010.
- 5** Solutions Management, "La nube: oportunidades y retos para los integrantes de la cadena de valor," pp. 2–40, 2012.
- 6** Secretaría de Educación Superior, Ciencia, Tecnología e Innovación, <http://www.educacionsuperior.gob.ec/>
- 7** Global e-Sustainability Initiative GeSI, <http://gesi.org/files/Reports/Smart%202020%20report%20in%20Spanish.pdf>.
- 8** Unión Internacional de Telecomunicaciones ITU and Climate Change, <https://www.itu.int/itu-news/manager/display.asp?lang=es&year=2008&issue=05&ipage=37&ext=html>
- 9** Real Academia Española RAE, <http://www.rae.es/>
- 10** Productividad en el aula - Reflexión Académica, http://fido.palermo.edu/servicios_dyc/publicacionesdc/archivos/269_libro.pdf.
- 11** A. Urueña and A. Ferrari, "Cloud Computing," p. 341; Creación:2012; Recuperado: 24 mayo 2016, 2012.
- 12** Microsoft Private Cloud Fast Track Reference Architecture Guide 2016, <http://fasttrack.microsoft.com/es-ES/>

Juglar Web: Módulo de programación visual y motor de visualización para la creación de multimedia educativa interactiva y el desarrollo del pensamiento computacional

José Enrique Álvarez Estrada¹, Nancy Aguas García², Carlos Enrique Hoy Chi³, Daniel Maximiliano Dzul Tun⁴

Universidad del Caribe, Coordinación de Ingeniería en Telemática,
Fraccionamiento, Tabachines, CP 77528 Cancún, Q.R.,
México

¹jeae@ucaribe.edu.mx, ²naguas@ucaribe.edu.mx, ³120300161@ucaribe.edu.mx,
⁴120300076@ucaribe.edu.mx

Resumen: En la actualidad existen herramientas de software que permiten la integración de Multimedia Educativa, es decir materiales didácticos que orienten y regulen el proceso de enseñanza-aprendizaje de estudiantes mediante la combinación de texto, color, gráficas, animaciones, vídeo, sonido, en un mismo entorno. Pero sólo un número reducido de ellas permiten además desarrollar el pensamiento computacional mediante programación visual, y la mayoría requieren del usuario conocimientos avanzados e invertir abundante tiempo y dinero para su uso y aprendizaje así como entornos poco amigables y fáciles de utilizar. Basado en un proceso de innovación y de desarrollo usando una metodología ágil, se propone una herramienta web de autoría denominada *Juglar Web* compuesta por un módulo de programación visual basado en bloques (BUI) y un motor de visualización para la creación de contenido interactivo multimedia. La herramienta desarrollada cumple con estándares de usabilidad y permite al usuario utilizar medios hechos por terceros que en suma hacen de Juglar una herramienta sencilla e intuitiva.

Palabras Clave: Multimedia, Multimedia Interactiva, Multimedia Educativa, Motor de Visualización, Programación visual (BUI).

Abstract: Nowadays there are software tools that allow the integration of Educational Multimedia: didactic materials that orient and regulate the teaching-learning process of students through the combination of text, color, graphics, animations, video and sound in the same environment. But only a small number of them also allow the development of computational thinking through visual programming, and most require the user advanced knowledge and invest abundant time and money for use and learning as well as unfriendly environments and easy to use. Based on a process of innovation and development using an agile methodology, we propose a web authoring tool called *Juglar Web* consisting of a visual programming module based on blocks (BUI) and a visualization engine for the creation of interactive multimedia content. The tool developed complies with usability standards and allows the user to use materials made by others that in brief make Juglar a simple and intuitive tool.

Keywords Multimedia, Interactive Multimedia, Educational Multimedia, Visualization Engine, Visual Programming (BUI).

Eje temático: Tecnologías en la innovación educativa e investigación.

1. Introducción

La educación tiene que mantenerse al día con las nuevas tendencias tecnológicas. Por ello, siempre debe buscar la mejora de los recursos, estrategias didácticas y materiales de apoyo, que resulten atractivos para los estudiantes.

Aunque existe una amplia variedad de herramientas que permiten a los docentes crear recursos didácticos multimedia, a menudo se les dificulta emplearlas debido a varios factores: 1) no son multiplataforma; 2) su costo es oneroso; 3) una vez adquiridas, su aprendizaje requiere una importante inversión de tiempo y esfuerzo.

En buena medida ello se debe a que son herramientas muy complejas, que incluyen funcionalidad tanto para la creación de medios (audio, video, animación) como para el proceso de autoría (*authoring*, integración de medios) propiamente dicho.

Debido a la gran abundancia de medios (imágenes *raster* y vectoriales, animaciones, audio, música, vídeo, etc.) disponibles en la red (YouTube, Instagram, Facebook, etc.), a los esquemas flexibles de licenciamiento que presentan como Creative Commons, pero sobre todo a la complejidad de dominar un software especial para crear cada uno de ellos, la mayoría de los maestros prefiere integrar medios ya existentes, hechos por terceros, a crearlos ellos mismos.

Así, se identifica claramente la necesidad de crear una herramienta que responda a esta demanda, a saber: integrar de la manera más sencilla posible (principio *KISS*, *Keep It Simple Stupid*) medios que ya se encuentren disponibles en La Red, y permitir su despliegue también a través de La Red (presumiblemente mediante un navegador).

2. Antecedentes

Algunas escuelas están integrando el pensamiento computacional desde el nivel preescolar. Como en Singapur, por ejemplo, donde a través del uso de tabletas y equipos de cómputo se brinda a los discentes funciones interactivas, multimedia, juegos y otras actividades que les ayudan a abordar la resolución de problemas y el pensamiento creativo (COSN, 2017).

Un docente de una escuela de secundaria de España descubrió que cuando sus estudiantes creaban vídeos, también exploraban aspectos del lenguaje relacionados a sus pasiones e intereses personales. Los estudiantes se comprometieron en el proceso completo de producción incluyendo la narración, el diseño, la grabación y la edición de vídeo (usando el programa *Camtasia Studio*) así como en la investigación y el licenciamiento del contenido, la corrección y la publicación final de sus productos (Miller y TechSmith, 2015).

Repensar los roles de los maestros implica mejorar su capacitación y desarrollo profesional. Si bien muchas aulas incorporan dispositivos digitales e instrucción diferenciada, como parte de un movimiento hacia el reconocimiento de que el aprendizaje no está fijado a un ritmo, lugar o tiempo específico, el desarrollo profesional a menudo todavía se ofrece en un formato único para todos. Sin embargo, el distrito de las Escuelas Públicas de Bismarck en Dakota del Norte está sentando un precedente: recientemente reunió a 35 maestros en siete escuelas secundarias y preparatorias, para crear opciones personalizadas de desarrollo profesional que

satisfagan las diferentes necesidades de sus maestros. Desarrollaron un entorno de aprendizaje asíncrono con multimedia interactiva, foros de discusión en línea y módulos de aprendizaje a su propio ritmo, que aprovechan la retroalimentación automatizada (COSN, 2017).

3. Materiales y métodos

3.1. Metodología Scrum

Para la construcción del proyecto se estableció el uso de la metodología ágil para el desarrollo de software *Scrum*, empleada para guiar las actividades de desarrollo, a saber: recolección de requerimientos, análisis, diseño, evolución y entrega. Dentro de cada actividad estructural, las tareas de trabajo ocurren de acuerdo a un patrón de proceso llamado *sprint* (Pressman, 2010, p.69).

En *Scrum* se realizan entregas parciales y regulares del producto final. Por ello, está especialmente indicado para proyectos en entornos complejos, donde se necesita obtener resultados pronto, donde los requisitos son cambiantes o poco definidos, donde la innovación, la competitividad, la flexibilidad y la productividad son fundamentales.

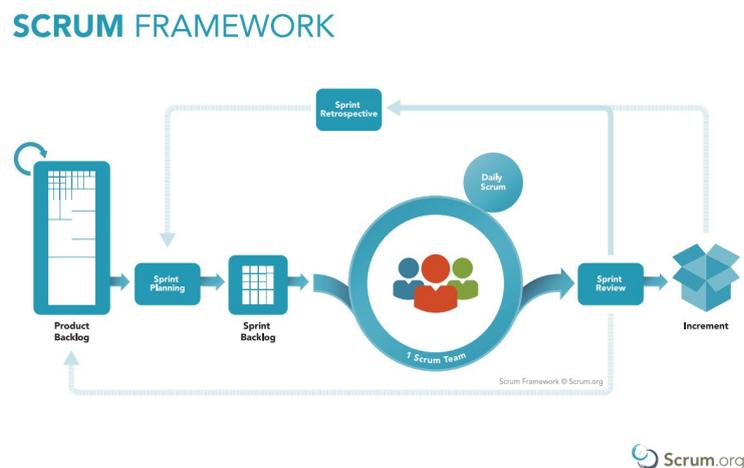


Fig. 1. Metodología Scrum.

Entre los elementos importantes de *Scrum* se encuentran los eventos y los artefactos. Los primeros están formados por el *sprint* y el *scrum diario*, siendo un *sprint* un bloque de tiempo de un mes o menos, durante el cual se crea un incremento de producto "terminado", utilizable y potencialmente desplegable. Por su parte, un *scrum diario* es un plan de trabajo para el equipo de desarrollo, que se utilizará durante las siguientes 24 horas.

Por su parte, los artefactos están conformados por el product backlog, una lista

ordenada de todo lo que podría necesitarse en el producto (actividades), y es la única fuente de requisitos para cualquier cambio a realizarse en el producto; y el sprint backlog, el conjunto de elementos de la lista de producto seleccionados para el sprint.

3.2. Metodología para la evaluación

Cuando se desarrolla un software, éste no debe únicamente cubrir las tareas para las que fue diseñado, sino que debe tener especial cuidado en el diseño de su interfaz, pues su aspecto puede resultar crucial para su aceptación o rechazo por el cliente. En *Juglar Web* es esencial su usabilidad: que permita integrar elementos multimedia de manera simple, con una rápida adaptación de la interfaz de usuario y facilidad de uso.

La aplicación de métodos que evalúen el nivel de usabilidad que tiene un sistema, ayuda a la búsqueda de la aceptación y éxito de un software, que dependa de que cumpla tanto con las necesidades como con las facilidades que requiere un usuario para interactuar con él.

Para realizar las pruebas de satisfacción a usuarios, se propone el uso del cuestionario *SUS (System Usability Scale)*, una herramienta metodológica muy similar a la Escala de Likert, pero especializada en medir la usabilidad de un objeto, dispositivo o aplicación. El cuestionario *SUS* fue desarrollado en 1986 por John Brooke como parte de la introducción de la ingeniería de usabilidad a los sistemas de oficina de Digital Equipment Co. Ltd. Su propósito es proporcionar un test fácil de completar, con un número reducido de 10 preguntas, fáciles de puntuar. Se utiliza generalmente después de que un usuario ha tenido la oportunidad de utilizar un sistema.

Tabla 1. Cuestionario SUS.

Pregunta No	
1	Creo que me gustaría usar JUGLAR con frecuencia.
2	He encontrado que JUGLAR innecesariamente complejo.
3	Pienso que JUGLAR fue fácil de usar.
4	Creo que necesito el apoyo de una persona para poder utilizar JUGLAR.
5	Me pareció que las funciones de JUGLAR fueron bien integradas.
6	Pienso que JUGLAR es muy inconsistente.
7	Pienso que la mayoría de la gente aprendería a usar JUGLAR muy rápidamente.
8	Encuentro que JUGLAR es muy difícil de usar.
9	Me siento confiado al usar JUGLAR.
10	Necesité aprender muchas cosas antes de ser capaz de usar JUGLAR.

Para obtener el puntaje, se debe sumar los resultados promediados obtenidos de los cuestionarios realizados a usuarios, considerando lo siguiente: las preguntas impares (1, 3, 5, 7 y 9) tomarán el valor asignado por el usuario, y se le restará 1. Para las

preguntas pares (2, 4, 6, 8, 10), será de 5 menos el valor asignado por nuestros entrevistados. Una vez obtenido el número final, se lo multiplica por 2.5.

El resultado final dará un valor entre 0 y 100, el cual representa el grado de aceptación de la herramienta evaluada.

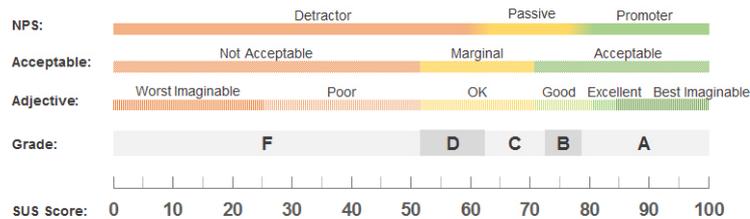


Fig. 2. Significado/Categoría puntuación SUS (Sauro J., 2018)

3.3. Selección de participantes

Para determinar el tamaño de la muestra, del grupo de prueba, se utiliza un modelo matemático desarrollado por Nielsen (1993), que modela el porcentaje de problemas encontrados en una evaluación de usabilidad para n usuarios, definida de la siguiente manera:

$$N(1 - (1 - L)^n)$$

Donde N es el número total de problemas de usabilidad en el diseño. Mientras que L es la proporción de problemas encontrados al evaluar a un usuario. El valor típico de L es del 31%.

Al trazar la curva para $L = 31\%$, se obtiene la gráfica de la figura 3. Se puede observar que se necesita probar con al menos 15 usuarios para descubrir todos los problemas de usabilidad. Nielsen menciona que, si se desea un solo número, se debe realizar la prueba a 5 usuarios, ya que permiten encontrar casi la misma cantidad de problemas de usabilidad que hallarían utilizando muchos más participantes en las pruebas.

Para estudios cuantitativos (con el objetivo de elaborar estadísticas), se debe probar con al menos a 20 usuarios para obtener números estadísticamente significativos (Nielsen J., 1993).

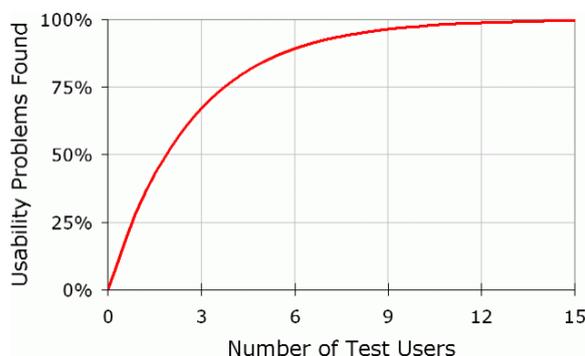


Fig. 3. Curva para L=31%

4. Diseño e Implementación

Como en toda herramienta se requiere levantar los requerimientos antes de comenzar con la fase de desarrollo (codificación). De lo mencionado hasta ahora queda claro que: 1) el usuario potencial de *Juglar Web* son profesores de todos los niveles educativos, que desean integrar (no elaborar) elementos multimedia para crear aplicaciones interactivas; 2) debe estar formado por dos módulos separados: un Editor de Historias (*StoryEditor*) y un Reproductor de Historias (*StoryTeller*); 3) ambos deben ejecutarse completamente en plataforma Web; 4) ambos deben ejecutarse independientemente el uno del otro, es decir, *StoryEditor* debe producir algún tipo de archivo (una *escaleta*, como dice la gente de radio y TV) que después *StoryTeller* reproduzca.

La figura 4 es el diagrama de casos de uso de *Juglar Web*; que ayuda a determinar las acciones que el usuario puede realizar. A partir de él se pueden obtener las historias de usuario, las cuales son una herramienta importante durante el levantamiento de requerimientos.



Fig. 4. Casos de uso Juglar Web (elaboración propia, 2019).

Los productos de autoría existentes se basan principalmente en dos metáforas de representación, a saber: el *carrusel de diapositivas* (i.e. Microsoft PowerPoint, LibreOffice Impress) y la *línea de tiempo* (i.e. Adobe Flash). El uso de la primera metáfora es ampliamente conocido y dominado por el público, pues el software de presentaciones de diapositivas forma parte de toda suite de productividad (i.e. Microsoft Office, LibreOffice); pero es una metáfora bastante restrictiva en su poder de expresividad. En cuanto a la segunda metáfora, la línea de tiempo, los maestros que se han capacitado en ella reportan que les resulta increíblemente difícil de emplear, sobre todo cuando se requiere animar un objeto a lo largo de la línea de tiempo: cuestiones como la sincronización cuadro por cuadro y la curva de trayectoria suelen resultarles muy difíciles.

Debido a que uno de los requerimientos ya mencionados de Juglar Web es cumplir con el principio KISS, se decidió hacer a un lado ambos enfoques y recurrir a un tercero: la metáfora es un escenario teatral, donde el director (i.e. el autor multimedia que emplea *StoryEditor*) decide qué acciones tendrán lugar paso a paso.

Como elemento innovador, se decidió implementar *StoryEditor* mediante una interfaz de usuario basada en bloques (BUI, *Block User Interface*). Las BUI se han popularizado recientemente gracias a aplicaciones como *Scratch* o *TurtleArt*, que emplean los bloques como elementos básicos (comandos, primitivas) para instruir a la computadora qué hacer. Google ha liberado *Blockly*, un marco de trabajo escrito en JavaScript para la creación de interfaces BUI, que se decidió emplear en *Juglar Web*.

Así, un bloque representará los distintos pasos (*steps*) que el director quiere que sucedan. El bloque de los pasos tendrá asociado un bloque de comando (es decir,

aquello que deberá ocurrir en ese paso en particular). Los comandos codifican acciones sencillas, similares a las que un director de teatro daría a su elenco y tramoyistas, a saber: aparecer una imagen (bloque *show*, bloque *display*), limpiar el escenario (bloque *clear*), cambiar los colores (bloque *color*), aparecer un personaje (bloque *sprite*). Pero, tal como si se tratase de una obra de teatro interactiva (donde el público juega un papel), también se codifican comandos como: definir secciones de la pantalla “sensibles” (bloque *button*); esperar a que el público (i.e. usuario) seleccione una de estas áreas interactivas sensibles (bloque *input*); y definir comportamientos variables (bloques *goto* y *gosub*) dependiendo de la selección (bloques *if*, *ifnot*). La Figura 5 muestra la forma en que se estructuraron dichos comandos mediante clases.

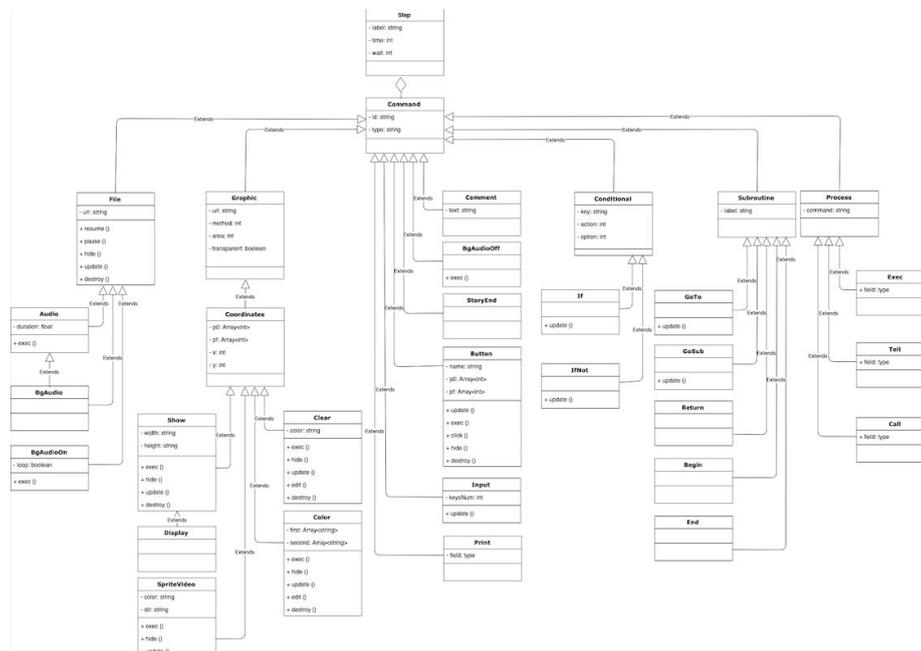


Fig. 5. Diagrama de clases de Juglar (elaboración propia, 2019).

Blockly soluciona el problema de crear la interfaz de usuario para la edición, pero aún resta el problema de reproducir multimedia en un navegador. Para abordarlo se recurrió a otro potente *framework*, llamado *Phaser*, que encapsula los principales comportamientos requeridos. En la figura 6 se presenta el ciclo de vida del objeto *Game* después de ser instanciado, y se observan los diferentes estados por los que puede pasar una historia durante su ejecución. El ciclo de vida de *Phaser* ejecuta los métodos *onPause()* y *onShutDown()* para detener las actividades actualmente en ejecución, y por tanto la historia que se está ejecutando, misma que puede reanudarse mediante el método *onResumed()*. El método *update()* se ejecuta constantemente, observando los cambios que se presentan en el objeto *Game*, y sirve para actualizar el escenario (objeto *Canvas*) sobre el cual trabaja *Phaser*.

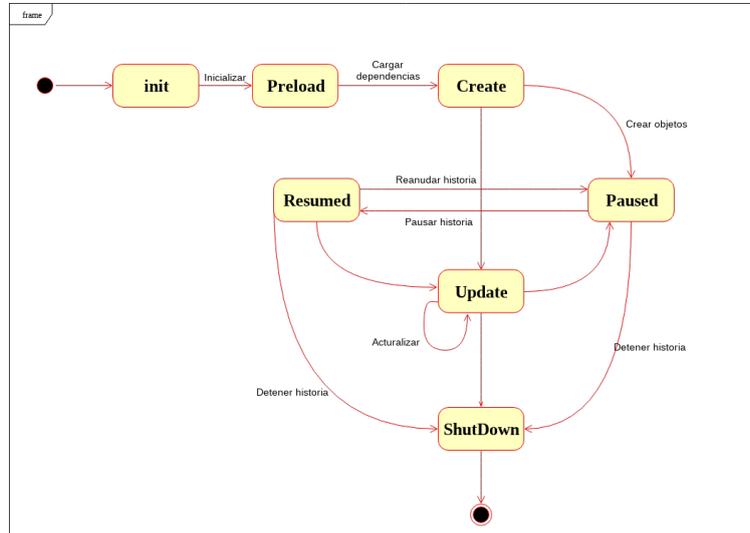


Fig. 6. Diagrama de estados Phaser.

Cabe señalar que Blockly proporciona una herramienta que también funciona con interfaz BUI, llamada *Block Factory*, que permite diseñar la BUI de la aplicación incluyendo el área que acepta bloques, la respuesta a los eventos (i.e. *Drag & Drop*) de los comandos de Juglar, etc. Block Factory gestiona el mapeo del bloque creado en el área de trabajo a un objeto JSON o JavaScript, que el propio Blockly puede leer posteriormente. También permite ubicar la posición de los parámetros del bloque, validar bloques de entrada que puede aceptar otro bloque, asignar colores, formas y textos de ayuda (*Tooltip*), etc. En resumen, es un excelente ejemplo de meta-interfaz de usuario.

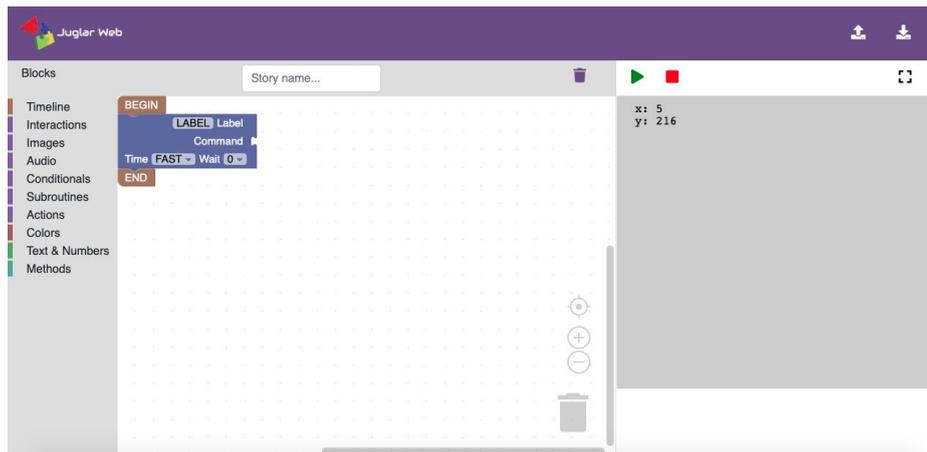


Fig. 7. Interfaz Juglar Web (elaboración propia, 2019),

En la figura 7 se observa la interfaz de usuario de *Juglar Web*, con la sección del listado de comandos en la parte izquierda y el escenario en la parte derecha. Dicho escenario representa a escala lo que se verá en el módulo de visualización *StoryTeller* cuando se reproduzca la historia creada en el área de trabajo del *StoryEditor*.

Es necesario implementar la funcionalidad de *StoryTeller*, iniciando con la detección de los bloques al momento de agregarlos al área de trabajo de *StoryEditor*.

En programación basada en componentes, existe una práctica conocida como *Global Bus Event*, utilizada para compartir mensajes entre componentes que no tienen relación alguna, pero necesitan realizar acciones o actualizar estados en el momento que ocurra cierto evento. Su funcionamiento es similar a la arquitectura *PUB/SUB* (*publish and subscribe*), en donde existen los *Publisher* que emiten eventos y notifican a las instancias llamadas *Subscribers*, que se encuentran escuchando a dichos eventos. En *Juglar Web* es necesario contar con comunicación en tiempo real, entre los módulos *StoryEditor* y *StoryTeller* al momento de crear las historias; es decir, se necesita que al momento de agregar un bloque de comando al área de trabajo (o modificar un atributo del mismo), el objeto que lo representa se cree o modifique en consonancia. Es decir, se necesita comunicación bidireccional entre *StoryEditor* (que emplea *Blockly* como *framework*) y *StoryTeller* (que emplea *Phaser*).

Mediante *Global BUS Event* se definen las variables que escuchan los eventos, y que pueden ser varias instancias simultáneas escuchando cambios. Cuando se agrega o elimina algo del objeto al que está escuchando, se notifica a todos.

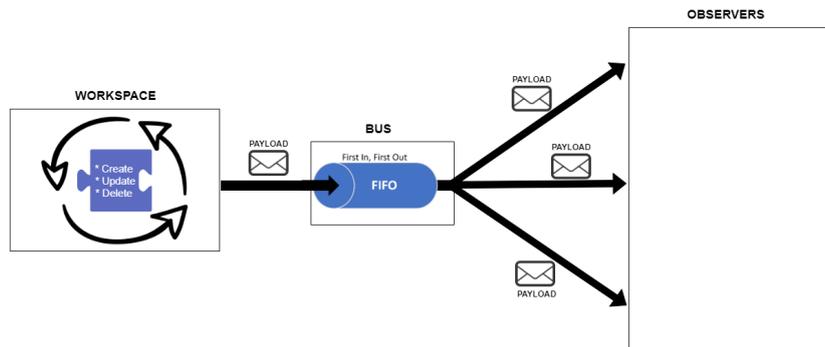


Fig. 8. Global BUS Event - Juglar (elaboración propia, 2019),

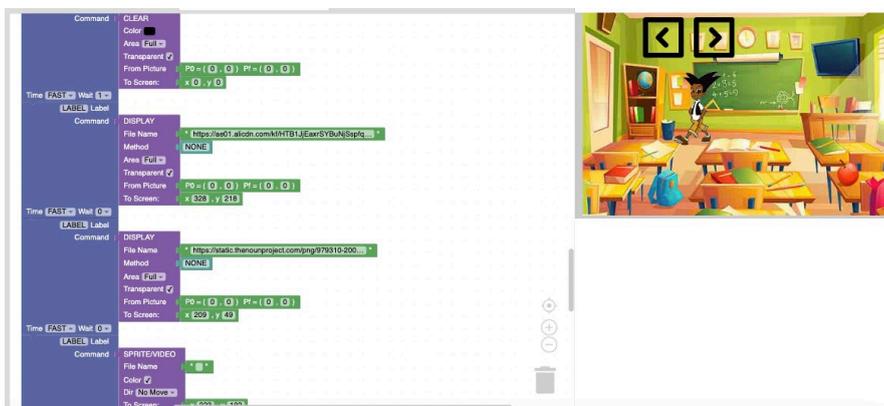
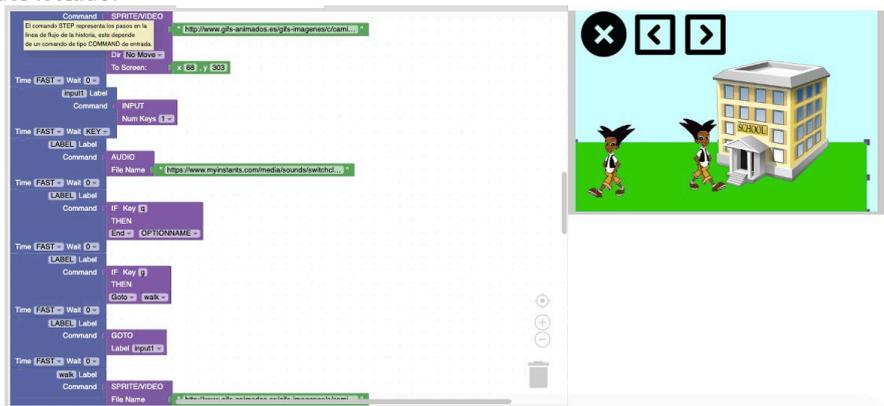
5. Resultados

Pruebas de Usabilidad

Una historia a evaluar está conformada por los siguientes bloques: paso (*step*), despliegue (*display*, presentar una imagen), control de flujo (*goto* y *gosub*, saltar de un paso a otro), animación (*sprite*, desplazar una imagen animada de un punto a otro),

áreas sensibles (*button*, seleccionar partes del escenario) y los comandos *clear* y *color* para cambiar colores, como haría un director mediante la iluminación.

La historia que se elaboró trata de un niño (*sprite*) que va camino a la escuela: se observa una imagen de un edificio escolar en la parte derecha del área de ejecución; tiene tres botones para interacción, uno para finalizar la historia y los otros dos para mover el personaje hacia la izquierda o la derecha. La escena cambia cuando se cumple la condición de que el personaje esté en la entrada de la escuela y se presione un botón x del teclado.



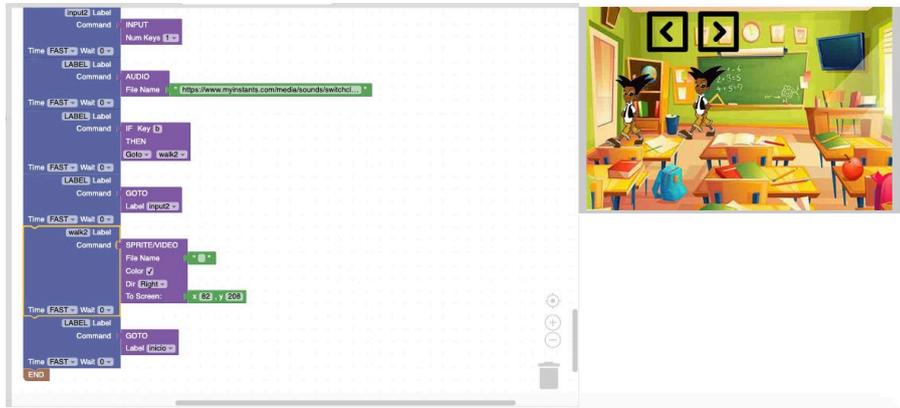


Fig. 9. Interfaz de Juglar (elaboración propia, 2019).

Desde el FLISOL (Festival Latinoamericano de Instalación de Software Libre) Cancún 2019, se presentó el proyecto *Juglar Web* a una audiencia de 24 participantes mediante el ejercicio de elaboración guiada de una breve historia. El objetivo era que los usuarios interactúan con Juglar y reportasen sus opiniones en un instrumento estandarizado, el ya citado *SUS* (*System Usability Scale*), desarrollado en 1986 por John Brooke, que consta de 10 preguntas, cada una con 5 opciones diferentes con valores de 1-5.

Tabla 2. Valores para cada una de las preguntas realizadas, a los usuarios evaluados.

	PREGUNTA										PONDERACIÓN	
	1	2	3	4	5	6	7	8	9	10		
USUARIO	1	5	1	5	1	4	1	5	1	5	1	97,5
	2	5	2	3	3	4	2	3	2	4	4	65
	3	2	1	3	5	3	1	2	3	3	3	50
	4	3	2	5	1	4	3	3	2	5	1	77,5
	5	5	4	4	2	3	2	4	2	5	2	72,5
	6	4	1	4	2	3	3	3	2	3	3	65
	7	5	1	4	3	4	2	5	2	5	1	85
	8	3	2	4	2	4	2	4	2	4	3	70
	9	5	2	4	3	5	1	4	1	4	1	85
	10	4	2	5	2	5	2	5	2	4	2	82,5
	11	3	3	2	4	5	3	3	2	4	3	55
	12	3	1	5	4	4	1	5	2	4	4	72,5
	13	3	2	4	2	4	3	5	1	3	3	70
	14	4	3	3	1	4	2	4	1	4	1	77,5
	15	4	2	4	4	4	3	5	1	4	1	75
	16	3	1	5	5	2	3	4	1	5	1	70
	17	4	3	4	3	3	2	4	2	4	3	65
	18	5	4	4	4	5	4	5	2	5	3	67,5
	19	5	1	5	1	5	5	5	2	5	1	87,5
	20	3	2	4	4	5	1	4	2	3	3	67,5
	21	4	3	5	1	5	1	5	1	5	1	92,5
	22	4	1	5	1	4	1	5	1	4	1	92,5
	23	4	1	5	1	4	4	5	1	5	1	87,5
	24	3	1	4	2	4	1	5	1	4	2	82,5
												75,52083333

En la parte derecha de la tabla, en la fila PONDERACIÓN se observa la ponderación de valores, tal y como se describió en apartados anteriores. El promedio da como resultado un valor SUS (celda color verde) que representa el grado de aceptación de la herramienta/sistema u objeto que se está evaluando. Para *Juglar Web* se obtuvo 75.5208333 en una escala entre 0-100, que se corresponde a un grado B (Bueno), y su valor de aceptación se encuentra entre un rango Aceptable.

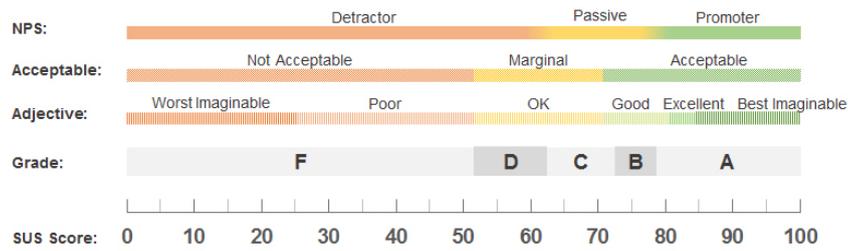


Fig. 10. Escala SUS.

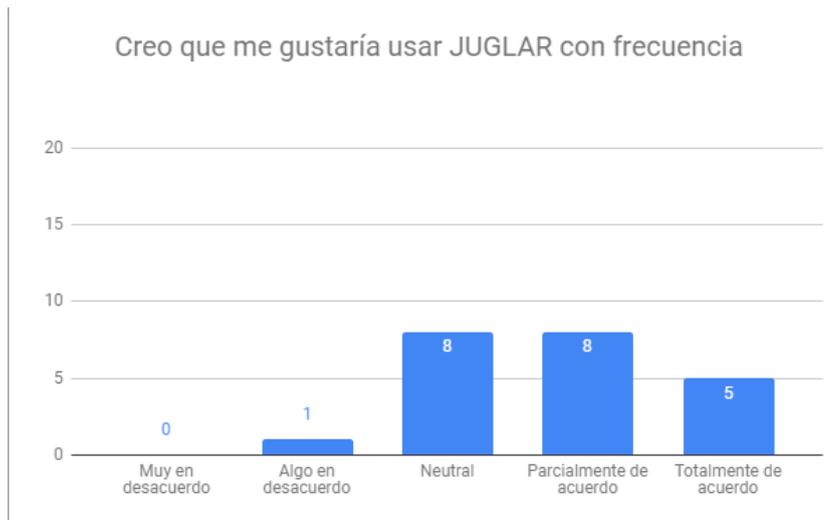


Fig. 11. Respuesta a pregunta de usabilidad. Elaboración propia, 2019.

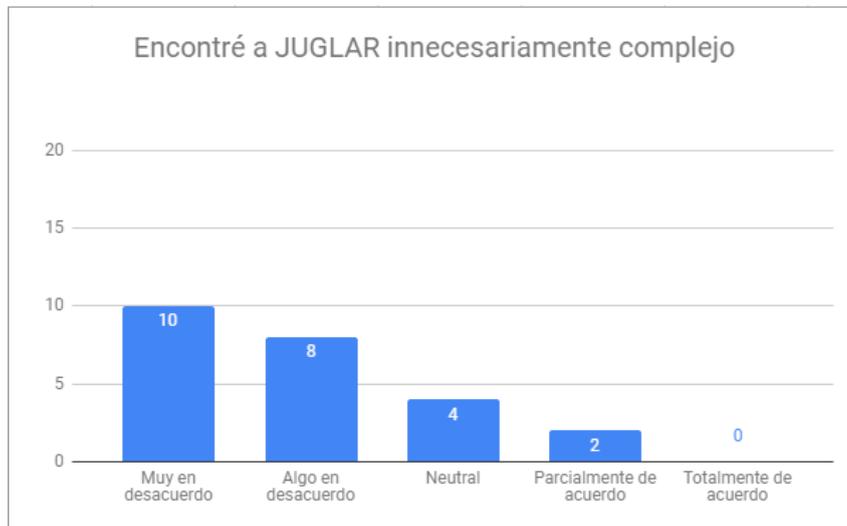


Fig. 12. Respuesta a pregunta de facilidad de interacción. Elaboración propia, 2019.

En la figura 12 se muestran resultados de la facilidad de usuario para interactuar con los diferentes componentes que conforman Juglar y de realizar una pequeña historia. Este resultado está directamente relacionado con la necesidad de que el usuario reciba apoyo de otra persona para poder utilizar Juglar.

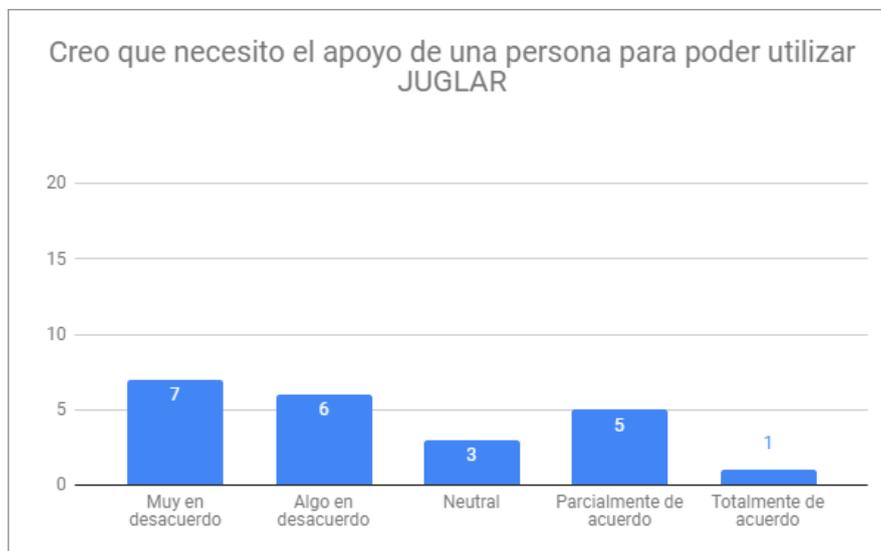


Fig. 13. Respuesta a pregunta de interfaz de usuario. Elaboración propia, 2019.

La interfaz de usuario de Juglar está pensada para mantener la simplicidad y la

facilidad de uso, desde los colores utilizados para representar cada una de las categorías de bloques, hasta el tipo de iconos que se utilizan, para la interacción con las Funcionalidades externas de Juglar (eliminar, cargar historia, exportar historia, etc.). De igual forma existe texto de apoyo para describir el funcionamiento de cada comando, facilitando de esta forma la familiarización entre usuario/herramienta.

6. Conclusiones

Los resultados indican que se ha logrado el objetivo en su mayoría, pero aún quedan cosas por mejorar. Durante la actividad guiada, los usuarios reportaron un grado de adaptación a Juglar muy bueno, a pesar de que varios de ellos manifestaron no haber utilizado alguna herramienta similar, que haga uso de bloques o de integración de medios.

A pesar de no haberse comparado los tiempos de desarrollo de *Juglar Web* con otras herramientas, de los resultados reportados por los usuarios durante las pruebas puede inferirse que su uso reducirá el tiempo dedicado a la autoría multimedia. La interfaz intuitiva basada en bloques del módulo de programación visual permite al usuario adaptarse rápidamente a ella, pues basta con arrastrar y conectar bloques para crear una historia en cuestión de minutos. Gracias al motor de visualización, el usuario puede observar cómo la historia va tomando forma con cada bloque que agrega al área de trabajo del módulo *StoryEditor*, tal como un director teatral lo hace durante los ensayos.

En esta versión de *Juglar Web* aún no se implementaron todos los comandos, ya que el alcance del proyecto se limitó a los comandos básicos de interacción, imágenes, audio, condicionales y subrutinas, pues con ellos es posible crear contenido multimedia básico. A pesar de ello y en base a los resultados de las pruebas, se puede concluir que *Juglar Web* es una herramienta que será bien aceptada gracias a su facilidad y comodidad de uso. Sin embargo, es necesario que se agreguen nuevas funcionalidades a futuro para poder disfrutar de una experiencia completa de la herramienta.

7. Referencias

1. Jakob Nielsen, Thomas K. Landauer. (April 29, 1993). A mathematical model of the finding of usability problems. CHI '93 Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems, 1, 206-213.
2. Álvarez, E. (2000). Juglar una herramienta GNU de autoría (Tesis de maestría). Fundación Arturo Rosenblueth para el avance de la Ciencia, México, D.F.
3. E. Pasternak, R. Fenichel and A. N. Marshall, "Tips for creating a block language with blockly," 2017 IEEE Blocks and Beyond Workshop (B&B), Raleigh, NC, USA, 2017, pp. 21-24.
4. Suárez Roldán, P. K. 2003. Animación y Visualización de Fenómenos Naturales. Tesis Licenciatura. Ingeniería en Sistemas Computacionales. Departamento de Ingeniería en Sistemas Computacionales, Escuela de Ingeniería, Universidad de las Américas, Puebla.
5. Pressman, R. (2010). Ingeniería del software. Un enfoque práctico. 7th ed.

1221 Avenue of the Americas, New York, NY 10020: The McGraw-Hill.

6. Instituto Nacional de Estadística y Geografía (México). (2015). Anuario estadístico y geográfico de Quintana Roo 2015. Recuperado 4 enero, 2019, de http://internet.contenidos.inegi.org.mx/contenidos/Productos/prod_serv/contenidos/espanol/bvinegi/productos/nueva_estruc/anuarios_2015/702825077228.pdf
7. Sergio Luján Mora (2001). Programación en Internet: Clientes Web (1ª edición). Editorial Club Universitario.
8. Sergio Luján Mora (2002). Programación de aplicaciones web: historia, principios básicos y clientes web (1ª edición). Editorial Club Universitario.
9. N. Fraser, "Ten things we've learned from Blockly" 2015 IEEE Blocks and Beyond Workshop (Blocks and Beyond), Atlanta, GA, 2015, pp. 49-50.
10. Douglas Crockford. (2018). How JavaScript Works. Virgule Solidus.
11. Lucy Black. (2012). Google Blockly - A Graphical Language with a Difference. 2019, de I programmer Sitio web: <https://www.i-programmer.info>
12. Phaser - HTML5 Game Framework (2018). Recuperado de <https://www.javascripting.com/view/phaser>
13. 2018,01. Concepto de Multimedia. Equipo de Redacción de Concepto Obtenido 2019,01, de <https://concepto.de/multimedia/>
14. Pablo. (2013). Imagen vectorial y raster. 2019, de Gruponex Sitio web: <http://www.gruponex.com>
15. Swift, Andrew. (1997), A brief Introduction to MIDI, SURPRISE (Imperial College of Science Technology and Medicine), consultado el 22 de agosto de 2012.

Capítulo 2

La conectividad y la infraestructura en la universidad digital

Modelo de clasificación de prioridad para la prestación de servicios de infraestructura tecnológica en la Universidad Estatal a Distancia

Rolando Rojas Coto ^a, Francisco Durán Montoya ^a

^a Dirección de Tecnología de Información y Comunicaciones, Universidad Estatal a Distancia, San José, Costa Rica
rorojas@uned.ac.cr, fduran@uned.ac.cr

Resumen. El presente documento describe una estrategia para entrega de servicios de TI a la comunidad universitaria, basada en un modelo de priorización, el cual se clasifica según su respectivo nivel de criticidad, para responder más eficientemente a las demandas tanto de la academia como de la administración, la cual fue implementada en la Universidad Estatal a Distancia de Costa Rica. Se presentan las ventajas obtenidas de este proceso y se describen los retos y desafíos que se presentaron durante el desarrollo del proyecto. Se presenta también el impacto en las diferentes áreas usuarias y el principal objetivo que consiste en permitir ofrecer un servicio diferenciado hacia el estudiante implementando una infraestructura convergente dinámica, escalable y modular, utilizando las últimas tendencias tecnológicas y acoplándolas al modelo universitario a distancia el cual es uno de los más consolidados de la región.

Palabras Clave: infraestructura tecnológica, servidores, convergente, virtualización, centro de datos, servicios de TI

Eje temático: La conectividad y la infraestructura en la universidad digital.

1. Introducción

La Universidad Estatal a Distancia (UNED) es una universidad pública costarricense cuya misión consiste en brindar opciones de educación superior a los diversos sectores de la población nacional “especialmente a aquellos que, por razones económicas, sociales, geográficas, culturales, etarias, de discapacidad o de género, requieren oportunidades para una inserción real y equitativa en la sociedad” [1].

Para lograr esta misión, la UNED hace uso de diversos medios y plataformas tecnológicas que permiten implementar un modelo de educación a distancia, el cual está centrado en el estudiante, razón por la cual, los medios deben permitir una interactividad y promover el autoaprendizaje, no sin antes olvidar una formación humanista [1].

La UNED, como universidad líder en la región latinoamericana en educación superior a distancia, ha experimentado un gran crecimiento a través de los años, con respecto al uso de plataformas tecnológicas para la prestación de sus servicios de cara a toda la comunidad universitaria, compuesta principalmente por estudiantes, docentes, investigadores, así como al público en general.

Para la prestación de los servicios digitales, la UNED cuenta desde el 2015 y hasta la fecha, con una plataforma tecnológica bastante sólida. El Proyecto de Consolidación de Servidores implementado entre los años 2012 y 2014 basado en técnicas de virtualización de servidores, apoyados con servicios en la nube, ha tenido resultados positivos en cuanto a la reducción de equipamiento y aumento de servicios en línea, tanto administrativos como académicos [2].

No obstante, debido al crecimiento de los servicios que se implementan en el Centro de Datos principal de la UNED, aunado a los problemas a nivel general de obsolescencia tecnológica, se hace necesaria la opción de aplicar una estrategia más adecuada, que brinde mayores capacidades de alta disponibilidad y principalmente, que permita sostener las demandas de aprovisionamiento de servidores y sus características, para de esta manera brindar un servicio masivo a las diferentes aplicaciones.

2. La demanda de recursos tecnológicos en los servicios académicos y administrativos

El aumento de los servicios en sitios web, almacenamiento, alojamiento de aplicaciones, bases de datos, sistemas de encuestas, entre otros, ha requerido de una mayor cantidad de recursos, en cuanto a la cantidad de servidores disponibles, capacidades de memoria, poder de procesamiento, espacio de almacenamiento en disco y conectividad de red. Un ejemplo de la problemática que se enfrenta en el manejo y la gestión de estos recursos, es visible con el sitio web del Campus Virtual (*campusvirtual*) de la UNED.

Este *campusvirtual* es utilizado por estudiantes y académicos de la UNED y consiste en una plataforma en línea basada en tecnología Moodle, la cual ha venido en constante crecimiento y ha tenido variaciones en la constitución de su infraestructura tecnológica, con el fin de mejorar la disponibilidad y la velocidad del servicio. A lo largo de los años este servicio ha tenido un incremento sostenido en cuanto a la cantidad de estudiantes

y en lo que a entornos virtuales se refiere. Los entornos virtuales se definen como la cantidad de cursos que se configuran en línea en el entorno del Moodle para que sean utilizados por los estudiantes durante el cuatrimestre.

Cada entorno del *campusvirtual* posee documentos en diferentes formatos, foros académicos, entrega de tareas, enlaces de videos, materiales didácticos interactivos, servicio de correo electrónico para los estudiantes, pruebas escritas en línea, entre otros. Estas herramientas permiten que los estudiantes procesen la información de su curso y cada uno de estos elementos se multiplica por la cantidad de estudiantes conectados simultáneamente, lo cual genera una gran cantidad de hilos de procesamiento, utilización de memoria y almacenamiento.

En este mismo ejemplo, la figura 1 muestra la cantidad de estudiantes matriculados en la plataforma Moodle y la figura 2, contiene el dato de la cantidad de entornos virtuales creados en los últimos años y se puede ver una tendencia de crecimiento constante.

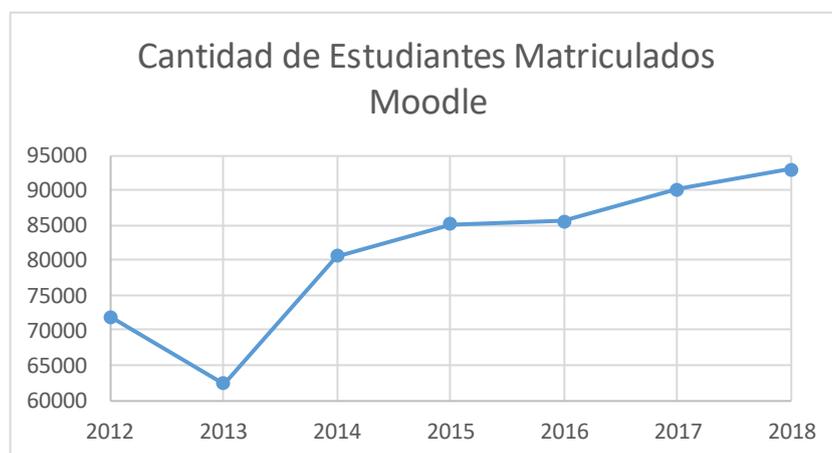


Fig. 1. Cantidad de estudiantes matriculados en Moodle desde el año 2012 hasta el 2018.

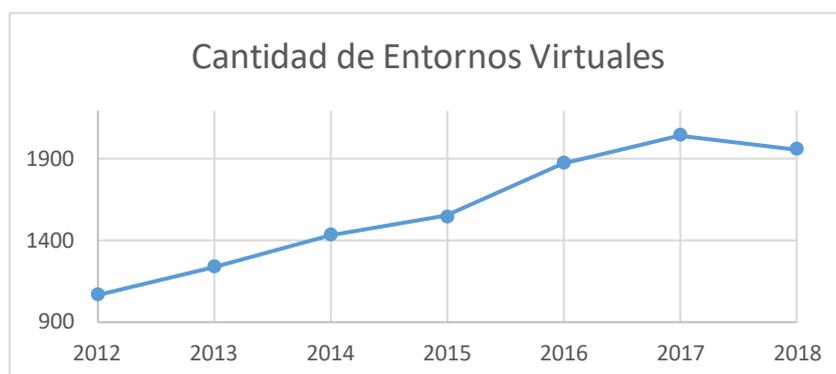


Fig. 2. Cantidad de entornos virtuales creados en Moodle desde el año 2012 hasta el 2018.

Estos datos de crecimiento además influyen en la manera en que los servidores se comportan, ya que al incrementar las conexiones concurrentes a la plataforma y principalmente las consultas a la base de datos se requiere por igual mayor procesamiento. Asimismo, existe una estrecha relación con respecto a la cantidad de almacenamiento empleado para el resguardo de la información del servicio del *campusvirtual* (Moodle). Ese incremento puede visualizarse también en términos de los terabytes (TB) consumidos y utilizados para el almacenamiento en disco requeridos para la restauración de entornos en caso de desastres. En el año 2012 este número apenas alcanzaba los 1.2 TB y durante el 2018 este número alcanzó los 5.8 TB.

La información aportada, a manera de ejemplo, corresponde también a una tendencia a nivel institucional ya que el crecimiento es similar con respecto a otros servicios de corte administrativo y académico, siendo la única constante el crecimiento para lo cual fue necesario modificar los modelos de clasificación de la prioridad en la prestación del servicio.

3. Modelo de distribución de servicios en clúster

El modelo de infraestructura tecnológica implementado en el año 2016, para la puesta en producción de los servicios académicos y administrativos de la UNED, estuvo constituido en su mayoría por elementos de virtualización, los cuales utilizan las herramientas de hipervisor de Hyper-V de Microsoft [2] a través de servidores configurados como clúster, y se representa según lo indicado en la figura 3.

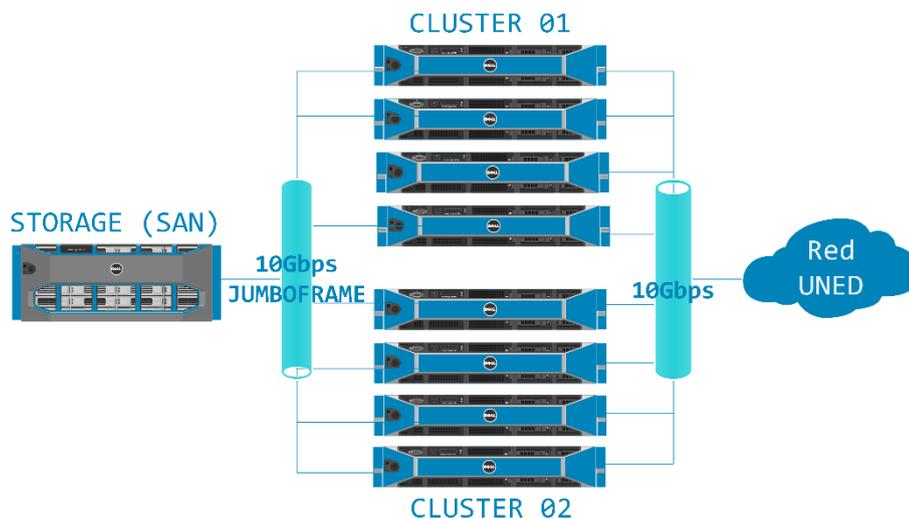


Fig. 3. Diagrama que representa la estructura de Clúster utilizada en el Centro de Datos principal de la UNED.

El equipamiento físico utilizado para este tipo de clúster estuvo conformado de la siguiente manera:

- Equipos servidores: DELL Power Edge R710, R820, HP Apollo 2800.
- Equipos de comunicaciones: Power Connect 6224.
- Sistema de Almacenamiento en Disco: Dell Equallogic 6100 y PS6110.

Este modelo se replicó en varios otros grupos de servidores que se mantenían trabajando para brindar servicios como: página web principal, intranet, *campusvirtual*, Entorno Estudiantil, Sistemas Académicos como tiempos académicos, formularios de becas; sistemas administrativos como Recursos Humanos, sistemas de vacaciones, librería virtual, entre otras aplicaciones en producción. Asimismo, los servidores a nivel de desarrollo también se alojan en este tipo de clúster para brindar acceso a los analistas programadores durante el quehacer diario. El escenario anterior se muestra en la figura 4 a continuación.



Fig. 4 Diagrama que representa algunos de los servicios distribuidos en los sistemas de clúster a partir de la estructura de virtualización.

A partir de este modelo se puede concluir que los servicios académicos, administrativos y de gestión propios del Centro de Datos, están distribuidos a lo largo de los clúster y servidores para virtualización. Esto conlleva un problema en cuanto a la gestión de los servicios, ya que los elementos de desarrollo se mezclan con los de producción y los servicios propios de infraestructura como DNS y Directorios Activos.

Se desprende de esta clasificación ad-hoc que realmente no se cuenta con una clasificación lógica de los servicios y principalmente de una base de infraestructura óptima para su ejecución, sino que se utiliza bajo la capacidad de la propia infraestructura. Cuando se necesitaba gestionar un mantenimiento del clúster, los servicios de baja prioridad permanecían mezclados con los de alta prioridad lo cual daba de baja servicios que debían estar lo más disponibles posible.

4. Implementación de un nuevo modelo de clasificación

Ante el incremento de los servicios administrativos y académicos, así como la demanda de recursos de servidores y almacenamiento, se buscaron fuentes de financiamiento para poder realizar la construcción de un nuevo Centro de Datos principal de la UNED en la Sede Central y la adquisición de infraestructura tecnológica que permitiera mejorar la prestación de los servicios. Ante esta y otras necesidades surgió el Plan de Mejoramiento Institucional (PMI) que unió varias iniciativas en materia de investigación y desarrollo.

El PMI se planteó como una respuesta a las necesidades y prioridades de la UNED en particular y del país en general, en las áreas de docencia, producción, investigación y extensión. Este plan enmarca acciones en el espacio específico dedicado a la modalidad de educación a distancia y dentro de los objetivos nacionales que se asignan al sistema de las universidades públicas, que son definidos a partir del entorno de políticas académicas establecidas en el Plan Nacional de la Educación Superior Universitaria Estatal (PLANES) y de políticas públicas determinadas tanto por el Plan Nacional de Desarrollo (PND) actual como por el reciente Plan Nacional de Ciencia, Tecnología e Innovación [3].

Gracias al PMI se han desarrollado diferentes iniciativas, varias de ellas orientadas al desarrollo de los servicios tecnológicos que posee la Universidad. Dentro de sus metas más importantes se destacan en materia de equipamiento y recursos tecnológicos las siguientes:

- Creación de un nuevo datacenter principal, que se encuentra ubicado en la sede central de la UNED, que atiende a toda la comunidad universitaria de la UNED [3]. Este datacenter contempla todos los requerimientos relacionados con las iniciativas propuestas en este proyecto y con todas aquellas que la UNED desarrolla de forma paralela.
- Desarrollo de la nube privada [3] de la UNED que permitiría el aprovisionamiento de nuevos servicios a estudiantes y funcionarios a futuro.
- Implementación de un datacenter alternativo [3] que se ubicará en el nuevo Centro Universitario de Cartago que será utilizado en caso de contingencias para dar continuidad en las operaciones en caso de caídas inesperadas o en caso de siniestros.

4.1 Implementación de Infraestructura Convergente

Los centros de datos convergentes se pueden considerar como la evolución de los centros de datos tradicionales. Su característica principal radica en que todos sus elementos se consolidan en un solo rack, habitualmente de un único proveedor. Todos los equipos y componentes que están en el rack son compatibles entre ellos y son manejados a través de una consola de administración centralizada, desde la cual se pueden gestionar todos sus componentes. El componente principal subyacente de esta tecnología es la virtualización la cual se explota en este tipo de infraestructura [4].

La arquitectura consta de diferentes elementos, tales como:

- *Capa de Almacenamiento*: En esta capa se localiza el almacenamiento de tipo SAN.

- *Capa de Red:* Está compuesta por switches con puertos CNA y pueden manejar redes de tipo LAN y SAN.
- *Capa de cómputo:* En esta se encuentran los servidores, es la base para la virtualización en la que se despliegan todos los servidores y aplicaciones del negocio.

Basado en este paradigma, se realizó la adquisición de equipamiento para sostener los servicios que requiere la UNED la cual utilizó el esquema de infraestructura convergente y que permitió obtener los beneficios que se detallan a continuación.

Disminución de la Obsolescencia Tecnológica. Con este equipamiento se logra migrar los servicios a nuevas plataformas físicas y lógicas de virtualización con novedosos elementos de monitoreo, gestión de plataforma, recursos actualizados. Se logra actualizar en su mayoría el equipamiento tanto de redes de datos como de servidores y almacenamiento, reemplazando equipos que estaban con una obsolescencia importante.

Aumento en la Capacidad de Recursos. La adopción de infraestructura convergente modificó la forma en que se configuraban anteriormente los servidores virtuales, ya que, al mantener toda la infraestructura en bloques, las capacidades de duplicación a nivel de almacenamiento utilizando disco de estado sólido, asimismo, los elementos como el manejo de hilos de procesamiento, y el aumento considerable en la capa de red de datos donde se aumentó entre 10 a 40 veces las velocidades de transmisión entre servidores y hacia los equipos de CORE.

Disponibilidad de los Servicios. Entre los años 2016 y 2017, uno de los servicios de misión crítica de la Universidad es el *campusvirtual*, había tenido una disponibilidad del 97,3% contabilizando aproximadamente 200 horas de caídas del servicio. Una vez que el servicio fue migrado a la infraestructura convergente, la disponibilidad aumentó a un 99,8% con un 0,1% de tiempo por actividades no planificadas (aproximadamente 8 horas). Este promedio ha mejorado la percepción de los estudiantes de cara a sus labores diarias en la plataforma.

Escalabilidad de la Infraestructura Convergente. El equipo convergente adquirido mantiene el potencial de crecimiento tanto horizontal como verticalmente, de forma que puede desarrollarse entorno a adquirir racks en bloques y también en la configuración actual, agregando mayor memoria, servidores tipo blade y almacenamiento. Además, que, en un entorno de convergencia, se simplifica la ruta para actualizaciones futuras de la plataforma y facilita el acoplamiento de tecnologías como aprovisionamiento en la nube e interconexiones desde Centros de Datos Alternos, esto siempre enfocado a la visión de alta disponibilidad para mejorar el rendimiento de las aplicaciones hacia el estudiante.

Capacidad de Autogestión. La infraestructura convergente actualmente adquirida, tiene la capacidad de ser monitoreada desde cualquier tipo de software de monitoreo, de igual manera posee una tecnología llamada "Call Home" la cual, ante cualquier evento de daño de alguna parte, emite una llamada a través de Internet a la fábrica y ésta realiza el despacho de la parte de manera inmediata hacia las oficinas donde se encuentra el Centro de Datos de la UNED. Otro elemento importante de autogestión del equipo es la capacidad de migrar las máquinas virtuales de los servicios que están siendo utilizados de manera masiva; por ejemplo, en tiempos de procesos críticos como matrícula, el equipo de manera automática logra migrar de discos normales a discos de

estado sólido de alta velocidad, esto para mejorar su rendimiento de cara al usuario final.

A manera de resumen se pueden esquematizar los componentes de la propuesta implementada, así como los beneficios percibidos, en la figura 5.

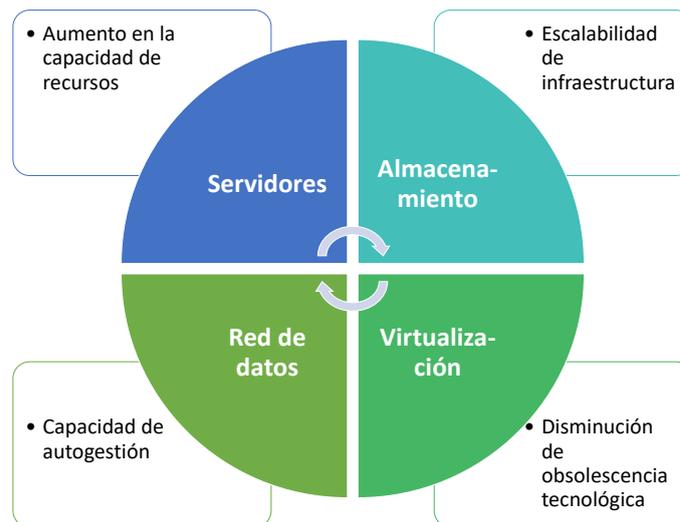


Fig. 5. Diagrama que representa la infraestructura convergente y los beneficios que han generado para la Universidad.

4.2 Modelo de priorización de servicios académicos y administrativos

La clasificación de los servicios por prioridad nace de la necesidad de la Administración de oficializar la atención de los servicios y darle las mejores condiciones a estudiantes y funcionarios.

La continuidad de las operaciones institucionales en materia de TI es una labor fundamental para brindar un adecuado soporte a los procesos sustantivos, así como para evitar posibles afectaciones a los usuarios. En ese sentido, la Institución debe desarrollar e implementar un modelo de continuidad razonable, basado en buenas prácticas y estándares. Se centra en los elementos básicos de la Continuidad de Operaciones, a saber: la identificación de objetivos y aplicaciones principales del negocio; el análisis de cada servicio basado en los criterios de urgencia e impacto para definir su prioridad; la definición de estrategias de control; el procedimiento para la atención de estas contingencias; los roles, responsabilidades y recursos; la comunicación de las medidas de continuidad; y finalmente la posibilidad de promover un cambio que permita aprender de las contingencias para generar una cultura más preparada a nivel organizacional [5].

Para el establecimiento de un plan de continuidad de negocios, las "Normas técnicas para la gestión y el control de las Tecnologías de Información (N-2-2007-CO-DFOE)" emitidas por la Contraloría General de la República, definen en la Norma 1.4.7

(Continuidad de los servicios de TI) lo siguiente:

La organización debe mantener una continuidad razonable de sus procesos y su interrupción no debe afectar significativamente a sus usuarios.

Como parte de ese esfuerzo debe documentar y poner en práctica, en forma efectiva y oportuna, las acciones preventivas y correctivas necesarias con base en los planes de mediano y largo plazo de la organización, la evaluación e impacto de los riesgos y la clasificación de sus recursos de TI según su criticidad. [6]

Esta norma, además resume las disposiciones generadas por la Contraloría General de la República y remitidas a la Institución en esta materia, específicamente en el informe N° DFOE-SOC-IF-08-2015 cuya disposición 4.13 solicita la elaboración de un "Plan de Continuidad de Negocios" y además detalla que el plan debe ser realizado desde el punto de vista técnico. Asimismo, indica que el Plan de Continuidad además de incluir actividades a desarrollar en caso de desastre, analiza las vulnerabilidades y genera acciones mitigadoras para dichas vulnerabilidades; por lo tanto, este Plan, no solo es un rehabilitador, sino que también desarrolla acciones preventivas con un mayor grado de planificación.

Dentro de este Plan de Continuidad de Servicios de TI, se establecen tres tipos de prioridades de atención, basado en un análisis de riesgo que utiliza como variables: el impacto y la probabilidad de caídas que puede presentar un servicio y según estos valores se conforman tres grupos de prioridad basados en: Alta, Media y Baja. El Plan utilizó como base este análisis de riesgo y categorizó los servicios brindados por el Centro de Datos en una de estas prioridades.

4.3 Clasificaciones Oro, Plata y Bronce para servicios de TI

Como se indicó anteriormente, hace algunos años los servicios de TI de la UNED se estaban implementando con una distribución poco estratégica en los diferentes clústeres, razón por la cual se aprovechó la etapa final de la implementación de la nueva infraestructura convergente, y basándose en el Plan de Continuidad de Negocios de TI, se realizó una distribución utilizando los tres posibles Acuerdos de Nivel de Servicio, según cada nivel de prioridad.

Según la propuesta los niveles quedaron distribuidos: Misión Crítica (Oro) para los servicios de prioridad alta, Atención Preferente (Plata) para los servicios de prioridad media y Mejor Esfuerzo (Bronce) para los servicios de prioridad baja, correspondientemente. El detalle de cada clasificación se detalla a continuación:

Prioridad Alta o Categoría Oro: Los servicios de prioridad alta, deben ser considerados como servicios de misión crítica por lo que no pueden detenerse y requieren un alto grado de redundancia y alta disponibilidad. Asimismo, al presentar fallos, deben ser atendidos inmediatamente.

Prioridad Media o Categoría Plata: Estos servicios deben tener cierto grado de redundancia y disponibilidad, deben ser considerados servicios de atención preferente. Asimismo, al presentar fallos, deben ser atendidos a la brevedad posible, siempre que no haya incidentes de misión crítica.

Prioridad Baja o Categoría Bronce: Estos servicios no requieren redundancia o

alta disponibilidad. Asimismo, al presentar fallos deben ser atendidos de acuerdo a la capacidad de atención, y siempre que no haya incidentes de misión crítica o de atención preferente, por lo que se pueden tratar como de mejor esfuerzo.

La distribución de servicios según categoría se realizó en conjunto con la administración y el equipo de TI y se muestra según lo indicado en la figura 6.

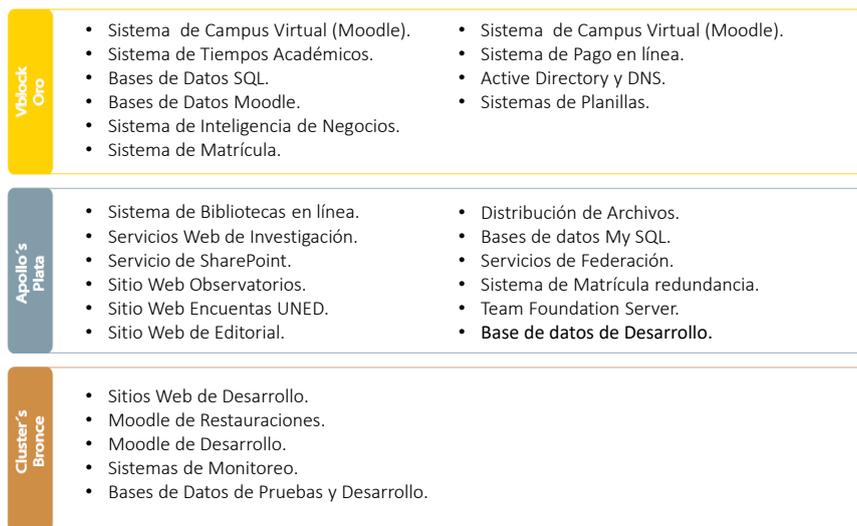


Fig. 6. Diagrama de clasificación de los servicios de TI por prioridad Oro, Plata y Bronce.

A continuación, se detallan los resultados obtenidos de los cambios realizados en el modelo de clasificación de servicios según la implementación final.

5. Resultados obtenidos y su impacto.

Los resultados obtenidos una vez completado este proyecto han sido muy satisfactorios, en cuanto a la transformación de la infraestructura hacia una tecnología convergente, la cual, no solamente evolucionó la forma de aprovisionar los servicios, sino que modificó la manera de trabajo a lo interno de la unidad encargada de la gestión de la infraestructura tecnológica.

Con relación a la implementación del modelo, a la fecha se ha logrado concretar la totalidad de los 57 servicios de TI acordados con la administración y se distribuyen de la siguiente manera: 20 servicios Oro (35%), 11 servicios Plata (19%) y 26 servicios bronce (46%), según se muestra en la figura 7.

PORCENTAJE DE SERVICIOS POR PRIORIDAD

■ Bronce ■ Plata ■ Oro

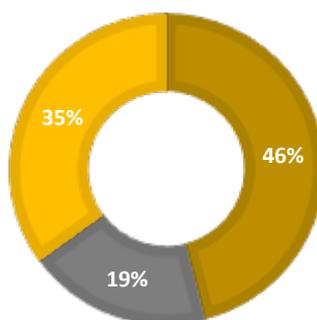


Fig. 7. Porcentaje de los servicios por prioridad Bronce, Plata y Oro.

Existen otros elementos que se destacan como resultados de impacto obtenidos en la implementación del proyecto, los cuales se detallan a continuación.

Gestión de Infraestructura Tecnológica. Anteriormente la infraestructura tecnológica que estaba implementada en el Centro de Datos principal de la UNED; requería de un soporte técnico y una atención diaria, múltiples fallas durante la semana laboral era una constante para el equipo humano de la Unidad encargada de Infraestructura. Ahora una vez terminado el proyecto, los equipos convergentes tienen la facilidad de autogestión, de igual manera la cantidad de incidencias disminuyó considerablemente, a tal punto que el recurso humano trabaja ya no en la resolución de problemas, sino que mayoritariamente en el mejoramiento y actualización de las plataformas en línea.

Reducción de Espacios del Centro de Datos. Con la entrada de la nueva unidad convergente se logra optimizar el espacio de los racks dentro del Centro de Datos, se eliminan equipos que están obsoletos, y en el futuro cuando se ejecute el escalamiento de la infraestructura tecnológica adquirida, se realizará en bloques los cuales están dentro de la misma infraestructura por medio de actualización de memoria, procesamiento con los servidores tipo blade y el aumento de almacenamiento en disco, lo cual permitirá el ahorro de espacio.

Infraestructura Tecnológica Homogénea. Con la entrada de los nuevos softwares hipervisores para virtualización, se aumenta la capacidad de gestión y la compatibilidad con los sistemas operativos hijos, los cuales anteriormente poseían problemas con algunas de las configuraciones de red, almacenamiento, también con el uso de las consolas de administración especialmente con sistemas de software abierto. Con el aprovisionamiento de los servidores en los nuevos sistemas de virtualización se logra

una infraestructura más homogénea con lo cual los servicios alojados tienen mejores condiciones y son propensos a menores fallos.

Percepción del Estudiante. Los servicios académicos que están dirigidos hacia los estudiantes aumentaron su disponibilidad, velocidad, la calidad del servicio y mejoró la experiencia del estudiante. La disminución de los problemas y solicitudes de soporte técnico en los periodos más críticos del cuatrimestre, aunado a la retroalimentación positiva de los comentarios y mensajes que se generan de las redes sociales oficiales de la Universidad, indican una mejoría en cuanto a la percepción del usuario final, elemento que se logró a partir de las adquisiciones realizadas con el Proyecto de Mejoramiento Institucional.

Visitas guiadas a la comunidad universitaria (Open House). Ante la puesta en marcha del nuevo Centro de Datos principal y la infraestructura convergente, se realizaron múltiples visitas y actividades para que la comunidad universitaria tuviera la oportunidad de conocer el nuevo equipamiento que se había adquirido y las nuevas instalaciones del recinto. Esto generó una experiencia invaluable para los estudiantes, funcionarios académicos y administrativos, al recibir charlas sobre el funcionamiento de los equipos, el ambiente donde se encuentran. La comunidad universitaria logró exponer sus necesidades como usuarios finales de las plataformas en línea, y se sensibilizaron sobre las acciones que se realizan cuando ocurren fallas en los servicios.

6. Conclusiones

Después de llevar a cabo el proyecto y la correspondiente implementación de prioridades de atención se pueden detallar algunos de los puntos importantes (aciertos) que se obtuvieron durante el desarrollo del mismo.

6.1 Aciertos

Utilización de Infraestructura Convergente La implementación de la infraestructura convergente ha sido uno de los puntos más importantes de la implementación ya que facilita la gestión y mantenimiento de la infraestructura y elimina puntos de fallo al reducir la cantidad de dispositivos y puntos de riesgo en la implementación. Los modelos de virtualización o de clústeres individuales no hubieran sido suficientes para implementar la estrategia planteada, ya que este tipo de infraestructura carece de modularidad.

Nuevo Centro de Datos. En el caso de los servicios universitarios de la UNED, no hubiera sido posible implementar un nuevo modelo sin contar con un Centro de Datos robusto y que permitiera la introducción de la infraestructura convergente nueva y del mejor aprovechamiento de los espacios actuales. Además, contar con la nueva infraestructura permitió pensar en modelos de mayor prestación para servicios de misión crítica mientras que con la infraestructura del Centro de Datos anterior no era posible.

Trasladado de Servicios: La arquitectura convergente adquirida, basada en la neutralidad tecnológica, fue pensada para permitir que, durante la implementación del modelo, se realizara un traslado de equipamiento al nuevo centro de datos

utilizando una metodología en caliente (transparente), donde se logra la cohesión tanto de hardware como software de diferentes marcas y tipos y sin impactar las operaciones de la UNED.

Plan Definido con Cronograma y Metas. En el caso del proceso de implementación, uno de los factores claves del éxito fue que desde las etapas de migración del centro de datos se contó con un cronograma y metas claras y realistas, donde el equipo de trabajo logró tener un plan definido desde el día cero. Este plan fue cumpliéndose y debido a que era realista en términos del tiempo, solamente requirió un seguimiento para gestionar que el mismo se llevara a cabo.

Prioridad del Proyecto. La administración de la UNED (principalmente representada por la Rectoría), asumió el proyecto y le dio alta prioridad al tratarse de un componente del plan del PMI, esto hizo que todas las metas se concretaran de una manera eficiente y eficaz. Este punto tiene un peso importante a lo largo de la implementación del proyecto, ya que hizo que las dependencias de la Universidad que estaban involucradas, actuaran de conformidad con el plan propuesto y se comprometieran con las diferentes etapas.

Como en todo proyecto de implementación durante la ejecución existieron algunos puntos que no se amalgamaron al plan y que es importante anotarlos para futuras referencias. Estos elementos se consideran desaciertos y se detallan en el siguiente apartado.

6.2 Desaciertos:

Mainframe. Uno de los problemas de la implementación del modelo es que en la UNED se cuenta con un mainframe ISERIES Power System, donde se encuentran implementados gran cantidad de aplicaciones y sistemas legados incluyendo una parte del ERP institucional. Debido a las particularidades de este sistema no es posible considerarlo como parte del nuevo modelo de implementación.

Plan de Sostenibilidad. Uno de los principales problemas que existen en la actualidad, es que a nivel institucional no se cuenta con un plan de sostenibilidad a mediano plazo, para garantizar la operación del nuevo Centro de Datos y del modelo de implementación. Durante la ejecución del proyecto, se solicitó este plan de manera oportuna con el fin de contar con una estrategia a nivel institucional que logre llevar a cabo una actualización y un escalamiento constante y definido, sin embargo, a la fecha no se ha logrado la oficialización de ningún plan de este tipo.

7. Impactos no esperados y desarrollos futuros

Existen algunos aspectos que no se contemplaron como parte de la estrategia y que actualmente generan varios riesgos o incluso impactos a nivel operativo. Además, existen otras mejoras que pueden llevarse a cabo a futuro según se detallan a continuación;

Complejidad Tecnológica. Aunque la infraestructura adquirida responde a un modelo de autogestión, se requiere de conocimiento especializado de la arquitectura y sus múltiples componentes para que el desempeño y rendimiento sean los correctos.

Además, la optimización de su característica principal que es la operación continua de los servicios, requiere personal altamente capacitado y sin los conocimientos adecuados no es posible sacarle mayor provecho al proyecto. Estos elementos podrían impedir el escalamiento uniforme de la infraestructura a través de los años.

Dependencia con Otras Unidades. Uno de los elementos no esperados de la gestión de todo el ambiente tecnológico del centro de datos, es que toda la operación mecánica, en donde se integran los componentes de: potencia, climatización, estructura física, sistemas contra incendios y de seguridad; son componentes que no son resorte de la gestión de TI, sino que son parte de las responsabilidades de otras Unidades y departamentos en la UNED. Por lo tanto, los mantenimientos, correcciones por garantías, modificaciones estructurales y de seguridad, dependen de terceros en donde la coordinación y ejecución de los trabajos se vuelve compleja y en muchos casos no se le da la importancia y atención requerida.

Centro de Datos Alterno. Como parte de los desarrollos futuros y de las iniciativas del PMI, se contempla la implementación en la UNED de un sitio alternativo a futuro, donde se podrán replicar los servicios de misión crítica de categoría "Oro". De tal manera que se debe contemplar todo lo relacionado con los modelos de recuperación ante desastres en este sitio.

Referencias

1. Información General, Universidad Estatal a Distancia, https://www.uned.ac.cr/images/documentos/2019/Informacion_General_2019.pdf
2. Durán, F, Rojas, R: Proyecto de Consolidación del Centro de Datos de la Universidad Estatal a Distancia de Costa Rica: Hacia una nube híbrida. Cuarta Conferencia de Directores de Tecnología de Información, TICAL2014. 134 a 148. 2014.
3. Plan de mejoramiento institucional. Universidad Estatal a Distancia. https://www.uned.ac.cr/images/ami/documentos/pmi_uned_agosto_2012.pdf
4. Santana, B. J. Estudio de factibilidad para la implementación de una infraestructura de hiperconvergencia de alta disponibilidad en la Data Center Experimental (Tesis de pregrado). Universidad de las Américas, Quito. 2019.
5. Plan de continuidad de Negocios de TI. Universidad Estatal a Distancia. 2015
6. Normas Técnicas para la Gestión y Control de las Tecnologías de Información (N-2-2007-CO-DFOE). Contraloría General de la República. <https://cgrfiles.cgr.go.cr/publico/docsweb/documentos/auditoria/normas-tecnicas-gestion-ti-n-2-2007-co-dfoe.doc>

Dockerización de una plataforma Moodle para migrarla hacia la nube

Milton Contreras^a, Carlos Ortega, Jairo Castañeda

^a Universidad Francisco de Paula Santander, Facultad de Ingeniería, Avenida Gran Colombia No. 12E-96 Barrio Colsag, San José de Cúcuta, Colombia
miltonjesusvc@ufps.edu.co, carlosivanoa@ufps.edu.co, jairoandrescp@ufps.edu.co

Resumen. Hace 5 años las tendencias y predicciones indicaba que para el 2020 la computación en la nube sería una tecnología establecida. Estas predicciones se cumplieron mucho antes y el ecosistema de educación virtual con Moodle no ha sido la excepción. Por lo anterior y buscando no quedar rezagados tecnológicamente, se planeó migrar una plataforma Moodle a la nube. Este trabajo describe una experiencia de migración a la nube de la plataforma Moodle del programa de Ingeniería de Sistemas de la Universidad Francisco de Paula Santander (UFPS). La migración tuvo en cuenta la arquitectura de despliegue, el proveedor de nube, los aspectos críticos en seguridad y el impacto en los usuarios. También, el monitoreo, la actualización de versiones y tecnologías, y la planeación del proceso de migración. El principal resultado fue una experiencia de migración exitosa y una mejora en la percepción de los usuarios. Se generó un plan de migración de la plataforma Moodle a la nube, enfocado en que el proceso se pueda replicar y que se consideren como factores de decisión: el rendimiento, la seguridad y los costos, así como la adaptabilidad a las necesidades y expectativas de los usuarios y la organización

Palabras Clave: Computación en la nube, Contenedores, Docker, Moodle.

Eje temático: La conectividad y la infraestructura en la universidad digital.

1. Problemática y contexto del proyecto

Históricamente, el Programa de Ingeniería de Sistemas ha liderado las transformaciones en TIC de la UFPS. Un ejemplo es la plataforma UVIRTUAL, creada desde el año 2003 y aún en operación con más de 15 años de funcionamiento resultado de un trabajo de grado de unos estudiantes para apoyarse sobre una plataforma e-learning que brindara asistencia al docente en el proceso de enseñanza., permitiendo una gran cantidad de beneficios como lo es:

Facilidad de la comunicación entre el estudiante y el profesor por medio de foros

Disponibilidad de material de estudio desde cualquier lugar

Automatización de las calificaciones

Potenciación para la enseñanza de programación con herramienta y laboratorios virtuales

Integración de material gráfico y animado complementario

Sin embargo, mantener una infraestructura inhouse ha desencadenado diferentes problemas como lo son:

- Costos elevados de mantenimiento
- Infraestructura desactualizada
- Perdida de información documental técnica
- Servicio fuera de línea en horas no laborales
- Dificultades para integrar nuevos plugins a la herramienta
- Recursos de hardware desperdiciados
- Imposibilidad de automatización de procesos
- Quejas por parte del usuario

Debido a estas constantes incidencias y después realizar búsquedas sobre las tendencias entorno a TI se propuso realizar la migración de la actual plataforma a Cloud Computing.

Para garantizar el éxito del proyecto se considerarán dos paradigmas de la Disciplina de los Sistemas de Información (ISD, por sus siglas en inglés Information Systems Discipline): (i) el paradigma de las ciencias del comportamiento, que busca desarrollar teorías para explicar y predecir el comportamiento humano y organizacional en contextos mediados por las TIC (búsqueda de la verdad) y (ii) el paradigma del diseño de la ciencia, que busca ampliar las capacidades humanas y organizacionales a través de la creación o modificación de artefactos apoyados en las TIC (búsqueda de la utilidad)[1]

Entendiendo que la metodología es el proceso que sigue el investigador o el ingeniero para conseguir sus objetivos. Dicho proceso debe ser rigurosos pero flexible y debe ser explícito para garantizar la objetividad. Por lo tanto, aplicando los planteamientos del Marco de Investigación en ISD, para este proyecto se propone la siguiente metodología.

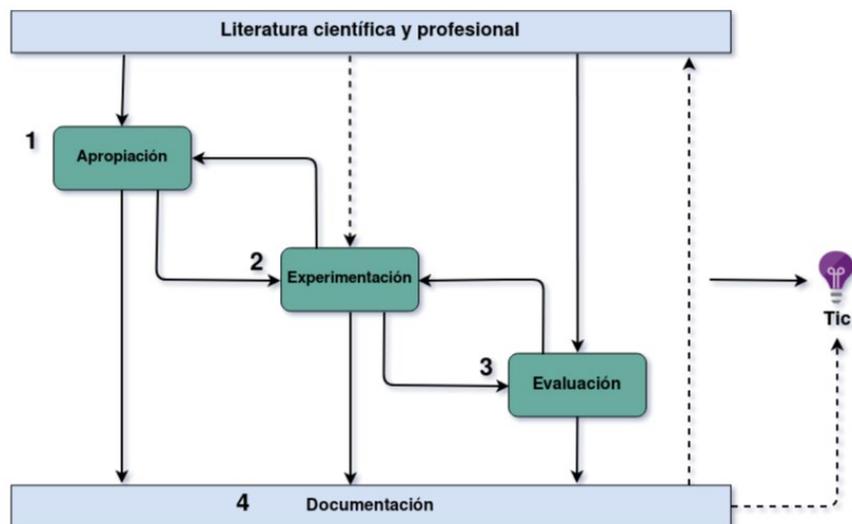


Fig. 1. Metodología aplicada para el desarrollo del proyecto Fuente: Elaboración propia

2. Descripción de la solución tecnológica implementada

Siguiendo el proceso metodológico propuesto, antes de realizar la migración a cloud computing de la plataforma LMS Moodle UVIRTUAL de la UFPS fue necesario responder tres preguntas:

1 - ¿Cuál es el estado del arte de Cloud Computing en el contexto educativo?

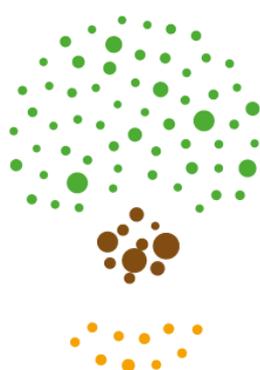
Para obtener el estado del arte o antecedentes del tema central del proyecto, se realizó una consulta en las bases de datos de Scopus y WoS, con las claves de búsquedas: "cloud computing + education". Los resultados se procesaron con la herramienta ToS (Tree of Science o Árbol de la Ciencia) de la Universidad Nacional de Manizales[2]

Para la clave de búsqueda "cloud computing + education" se obtuvieron 133 registros.

Los registros obtenidos de la búsqueda fueron exportados desde Web of Science e importados al Árbol de la Ciencia (Tree of Science). A continuación, se resumen el árbol y principales autores y tópicos:



TREE VIEW



●	Kalemis K, 2014, , V, P677
●	Workineh M, 2017, , V
●	Shana Z, 2017, INT J EMERG TECHNOL, V12, P168,
●	Eurich M, 2015, COMM COM INF SC, V554, P61,
●	Sultan N, 2010, INT J INFORM MANAGE, V30, P109,
●	Arpaci I, 2017, COMPUT HUM BEHAV, V70, P382,
●	Gonzalez-Martinez JA, 2015, COMPUT EDUC, V80, P132,
●	Ercan T, 2010, PROCD SOC BEHV, V2, P938,
●	Armbrust M, 2010, COMMUN ACM, V53, P50,

Fig. 2. Árbol de la ciencia [cloud computing + education]. Fuente: Elaboración propia usando la herramienta ToS.

De acuerdo a la búsqueda y la lectura preliminar (título, resumen, conclusiones) de la literatura, se identificaron algunos de los tópicos y autores principales asociados a este proyecto, lo cual se resume a continuación:

Tabla 1. Resumen de los resultados de cloud computing + education. Fuente: Elaboración propia

Posición	Referencia	Tema	Descripción
Raíz	[3]	Nube y beneficios de implementarla.	Define la nube desde diferentes vistas, y nombra los 10 principales obstáculos y oportunidades de la nube.
Raíz	[4]	Nube aplicada a la educación para lograr ventajas económicas.	Nombra la nube como una excelente alternativa para las universidades que están especialmente desprovistas de presupuesto para operar sus sistemas de información de manera efectiva y lo que se puede hacer para aumentar los beneficios de estudiantes y docentes.
Tronco	[5]	Nube aplicada a la educación para lograr ventajas económicas.	Expone a la cloud computing como una oportunidad para las universidades, para lograr reducir costos y algunos aspectos que impiden que lo adopten.

Hoja	[6]	Nube aplicada a la educación para lograr ventajas pedagógicas.	Muestra que se puede usar la computación en la nube como un proceso didáctico alternativo (bajo costo, aprendizaje personal, aprendizaje interactivo y el aprendizaje de muchos a muchos) en la educación primaria, secundaria y superior.
Tronco	[7]	Nube aplicada a la educación para lograr ventajas pedagógicas.	Revisión literaria sobre el uso e investigación de la computación en la nube en la educación. Además de identificar riesgos y sus escenarios para tener ventajas significativas.
Hoja	[8]	Nube aplicada a la educación para lograr ventajas económicas.	Lograr soluciones de HPC (high performance computing) implementando cloud computing en la educación. Apreciar la nube como un medio para lograr mayor transparencia y eficiencia en la educación.
Tronco	[9]	Nube aplicada a la educación para lograr ventajas pedagógicas.	Lograr gestión de conocimiento a través de la computación en la nube.
Hoja	[10]	Nube aplicada a la educación para lograr ventajas pedagógicas.	Aspectos de preparación y base para aplicar cloud computing en la educación superior en los Emiratos Árabes.
Hoja	[11]	Modelo de adopción de nube aplicada a las IES (Instituciones de educación superior).	Identifica las capacidades críticas y desarrolla un modelo para evaluar la preparación para la adopción de la informática en la nube de las instituciones de educación superior antes de iniciar la adopción.

Partiendo de la literatura encontrada de cloud computing, se evidencia que dentro de los tópicos de tendencia asociados en educación se identificó la arquitectura de containers Docker[12] Por lo tanto se decidió realizar una segunda búsqueda en la literatura.

Para la segunda búsqueda, se tuvo como clave “containers + linux”, inicialmente, se había usado la clave “containers + architecture”, que solo arrojaba 50 resultados y el árbol de la ciencia sugiere un rango entre cien (100) y quinientos (500) registros, con la finalidad de poder obtener un árbol consistente. Por lo tanto, se hizo el cambio en la clave de búsqueda a “containers”. Ésta segunda clave era muy general y arrojaba literatura sobre contenedores de carga de transporte marítimo, una consecuencia del uso de metáforas en computación[13] (el concepto de “containers” surge en la comunidad de software libre linux como metáfora de la cadena de suministros y logística de las empresas marítimas). Entonces, fue necesario refinar la búsqueda, quedando finalmente

la clave “containers + linux”, para un total de 94 registros.

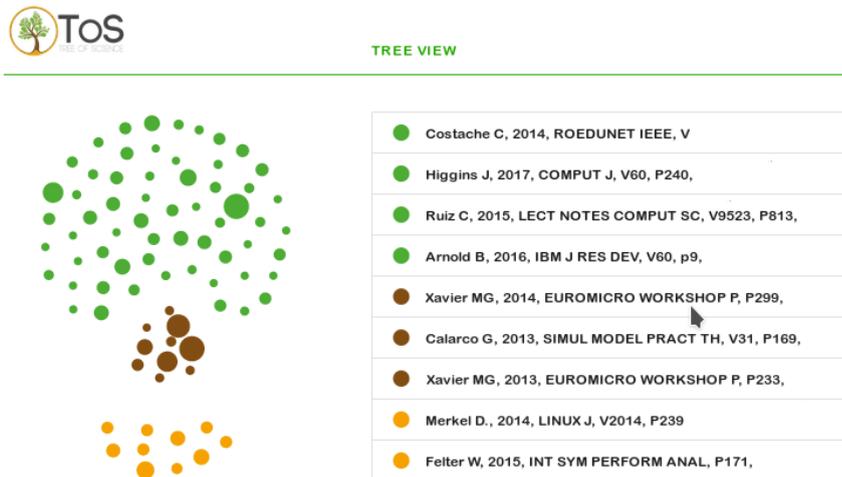


Fig. 3. Árbol de la ciencia [containers + linux]. Fuente: Elaboración propia usando la herramienta ToS.

De acuerdo a la búsqueda y la lectura preliminar (título, resumen, conclusiones) de la literatura, se identificaron algunos de los tópicos y autores principales asociados a este proyecto, lo cual se resume en la tabla.

Tabla 2. Resumen de los resultados de cloud containers + linux. Fuente: Elaboración propia

Posición	Referencia	Tema	Descripción
Tronco	[14]	Containers para virtualizar redes	Presenta una implementación de containers para una plataforma de redes heterogéneas, y nombra las limitaciones y sus posibles mejoras.
Tronco	[15]	Containers en el contexto de HPC	Realiza una serie de experimentos para realizar una evaluación del rendimiento de la virtualización basada en contenedores para HPC.
Raíz	[16]	Docker como tendencia y sus beneficios	Docker como una arquitectura de container y una solución para lo que él llama “infierno de las dependencias”, de una forma ligera, portátil y aislada.
Tronco	[17]	Ventajas de implementar containers	Realiza un experimento de los contenedores con aplicaciones de recursos intensivos (MapReduce. MR). Los resultados muestran cuál arquitectura container es la mejor en relación al rendimiento y su configuración.

Hoja	[18]	Containers como servicio	El documento presenta una solución para permitir el aprovisionamiento bajo demanda de contenedores Linux utilizando Software-Defined Networking, bajo un enfoque flexible para tratar incluso los recursos de nivel de control "como un Servicio".
Raíz	[19]	Ventajas de implementar containers	Comparan el rendimiento de una VM, con el uso de containers. Los resultados muestran un mejor o igual rendimiento de los contenedores en todas las pruebas.
Hoja	[20]	Containers en el contexto de HPC	Este documento evalúa el rendimiento de las soluciones de contenedor basadas en Linux aplicado en contextos de HPC (High performance Computing), Mostrando que la tecnología de contenedores ha madurado a lo largo de los años y que están en mejora continua.
Hoja	[21]	Containers como servicio	El desarrollo de IBM Containers, un servicio en la nube que proporciona a los desarrolladores las características de una nube PaaS y la flexibilidad de una nube IaaS, bajo una arquitectura de contenedores.
Hoja	[22]	Containers en el contexto de HPC	Implementa contenedores Docker en el entorno de clúster y evalúa su tiempo de ejecución para la ejecución paralela de alto rendimiento, enfocado en el flujo de usuarios de informática científica.

Actualmente la tecnología que lidera el mercado de containers es Docker el cual es un proyecto open source[23], el cual facilita el despliegue, migración y replicación a través de nube[24] muy utilizado por las empresas de tecnología[25]

2 - ¿Cuál es la nube más apropiada para realizar la migración y cuáles deberían ser los criterios para de selección de la nube?

Para el desarrollo del experimento de migración, se decidió tener en cuenta solo las 3 principales nubes públicas, basados en las nubes líderes según Gartner[26] las cuales son: Google Cloud, Amazon AWS y Microsoft Azure como se muestra en la siguiente gráfica.

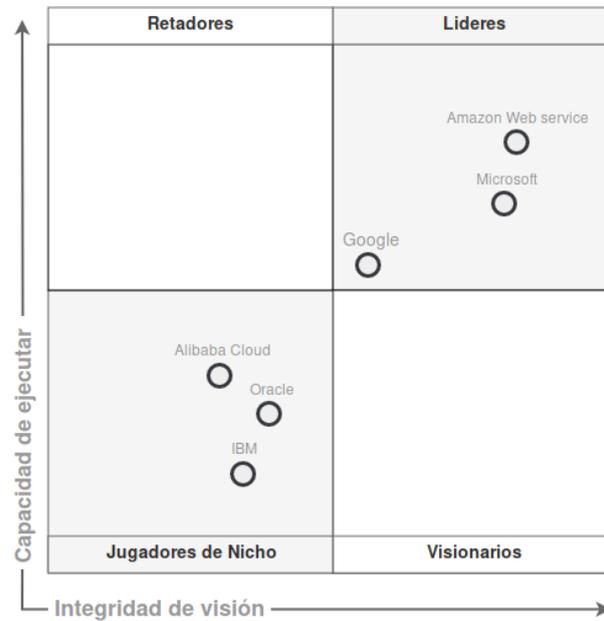


Fig. 4. Cuadrante de Gartner sobre las principales nubes . Fuente: Elaboración propia basada en la publicación original en inglés[26]

Teniendo como alternativa las nubes líderes anteriormente propuestas, se procedió a evaluarlas en 3 criterios básicos que fueron: economía, rendimiento y seguridad.

Criterio Económico

Para evaluar económicamente las nubes, se proyectó una infraestructura de referencia de 4GB RAM, 120GB SSD, 1 vCPU en especificación de recursos a un mes, sin embargo, si la nube no brindaba exactamente los recursos solicitados, se decidió seleccionar las especificaciones más cercanas ofrecidas por la nube.

Las consultas fueron realizadas a la fecha de 5 de febrero de 2019 y el resumen de la consulta se muestra en la siguiente tabla.

Tabla 3. Comparación de precios por nube. Fuente: Elaboración propia.

Nube	Características			Precio (USD)
	vCPU	RAM	SSD	
Google Cloud	1	4	120	34.13

Amazon AWS	2	4	120	49.34
Microsoft Azure	1	3.5	128	60.90

Criterio de Rendimiento

Las nubes líderes poseen muchos centros de datos repartidos en zonas estratégicas por todo el mundo con un total de 63 de Amazon AWS[27], 54 de Microsoft Azure[28] y 61 de Google Cloud[29]

Por tal motivo, es muy importante para el rendimiento, desplegar en el centro de datos con mejor latencia

Para evaluar el rendimiento de las nubes, se realizó una prueba a cada proveedor que consiste en medir la latencia desde el campus universitario hasta todos sus centros de datos, por lo cual fue necesario realizar (i) un listado de los centros de datos y (ii) realizar un test utilizando herramientas encontradas en la comunidad como gcping[30], azurespeed[31], cloudping[32]. El resumen de los resultados se presenta a continuación.

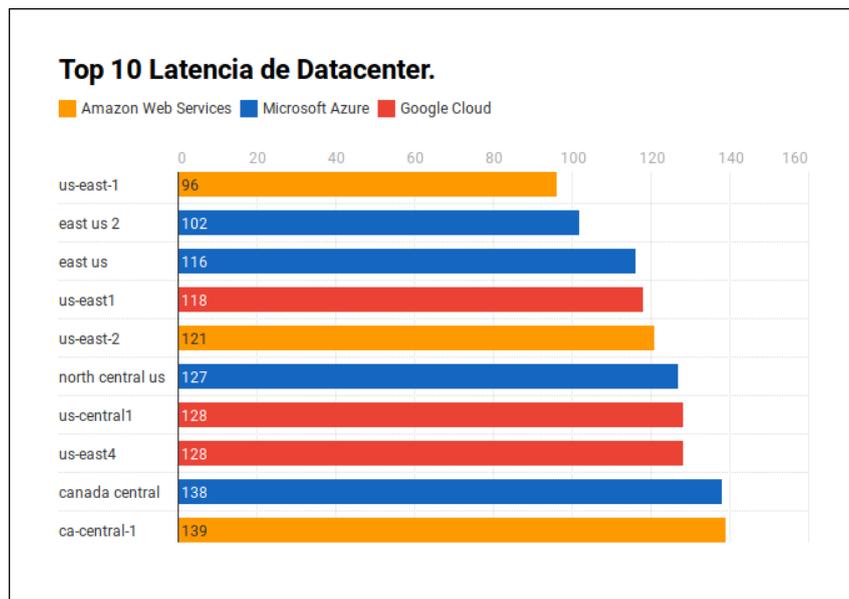


Fig. 5 Clasificación de los datacenter de menor latencia a mayor latencia de cada nube. Fuente: Elaboración propia.

La nube que tiene mejor rendimiento es Amazon AWS, seguida por Microsoft Azure y muy de cerca Google Cloud.

Sin embargo según el blog tech de Google la latencia no debe ser un valor único de selección[33], debido a que hay un valor promedio donde la percepción para el usuario

no es notable, y dado una diferencia menor de 22 (ms o puntos) entre los mejores registros de cada nube, fue necesario añadir otro criterio de rendimiento, el cual basados en la literatura, encontramos que la arquitectura implementada genera un impacto sobre el rendimiento[34]

3 - ¿Cuál es la arquitectura más apropiada para realizar la migración?

Arquitectura actual

Inicialmente, antes de la migración, UVIRTUAL estaba alojada en un servidor físico y dentro de él se tenía el servicio de bases de datos, el servicio de Apache, la API de PHP y la aplicación LMS Moodle, pero, en versiones deprecadas y obsoletas. Esta arquitectura base se presenta en la figura.

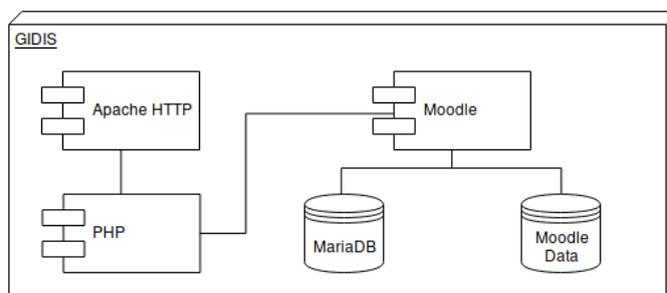


Fig. 6. Representación de la arquitectura actual de la uvirtual. Fuente: Elaboración propia.

Es necesario para seleccionar la arquitectura más apropiada, que previamente se actualicen las versiones de moodle y sus dependencias a la última versión estable, por lo tanto, para actualizar se siguió la estrategia de "lift and shift"[35] que consiste en hacer un rehospedaje de la aplicación ya que es más fácil optimizar sobre la nube. Esta estrategia se aplicó en la nube de Google.

Arquitecturas Propuestas

Luego se procedió a proponer 4 arquitecturas con la finalidad de ser evaluadas, estas arquitecturas son propuestas teniendo en cuenta el impacto de la implementación con dockers[36] y la utilización de varios niveles[34], Las arquitecturas propuestas se presentan a continuación.

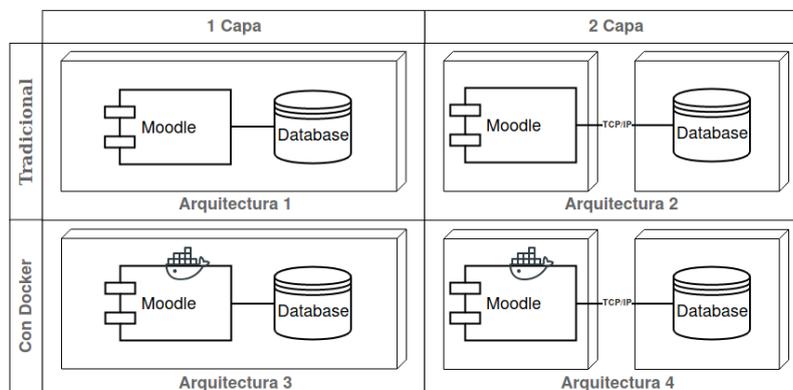


Fig. 7. Opciones de arquitectura propuestas para evaluar. Fuente: Elaboración propia

Arquitectura 1: Representa la arquitectura actual pero desplegada en nube, por lo tanto, no sigue las buenas prácticas de separación por capas y se espera que obtenga los resultados más deficientes.

Arquitectura 2: Aplica la separación por capas, desplegando la base de datos en otra instancia, por lo que se espera una mejora en los resultados y un aumento en los costos.

Arquitectura 3: Agiliza el despliegue implementando Docker, pero, en una sola capa, por lo que no se espera una mejora muy significativa, aunque, podría ser la mejor alternativa costo-beneficio.

Arquitectura 4: Emplea las buenas prácticas de separación por capas, además implementa Docker, por lo que se presume sea la arquitectura ideal, con el mejor puntaje en los resultados de las pruebas.

Evaluación de las arquitecturas

Para evaluar las arquitecturas se seleccionó el rendimiento, la seguridad y los costos como atributos a evaluar y para el diseño de las pruebas se siguió la metodología de test en nube propuesta por Martín Pol, Jeroen Mengerink, Kees Blokland[37].

Teniendo los atributos a evaluar y la metodología para medirlos, se procedió a realizar las pruebas.

Prueba de rendimiento

Pregunta: ¿Resiste la máxima capacidad de usuarios simultáneos?

Objetivo: Se realizará un test de carga usando el perfil del usuario visitante

Para realizar el test de carga, se utilizó la herramienta Jmeter la cual es una herramienta open source muy utilizada en industria de las pruebas

Para medir el consumo de recursos de procesador, memoria, almacenamiento y red, se implementaron herramientas de monitoreo, las cuales son un eje fundamental en la

nube[38] estas herramientas fueron las brindadas por el proveedor de nube, htop y monitorix. La diferencia de utilizar htop y monitorix, es que aunque ambas brindan datos de monitoreo, htop los muestra en tiempo real y por consola[39], mientras que monitorix[40] guarda un historial y posee un cliente web.

A continuación, se muestra los resultados de la prueba.

CPU

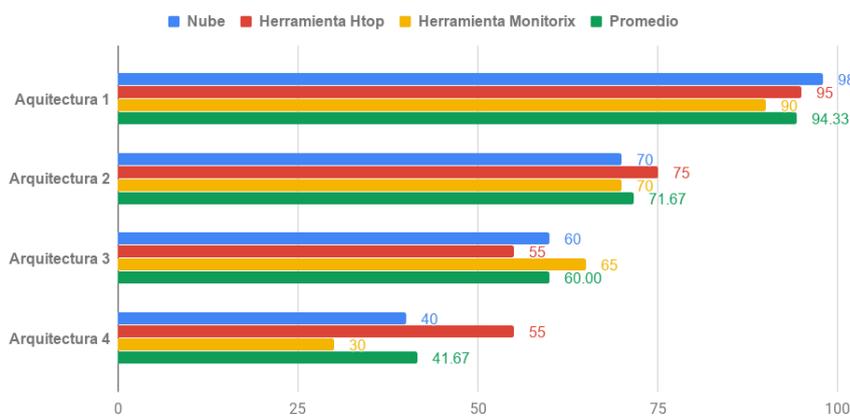


Fig. 8. Medición del uso del cpu en el eje y (las arquitecturas) en el eje x (el puntaje consumido en porcentajes sobre 100) medidos por tres herramientas, de color azul (nube), color rojo (herramienta htop) color amarillo (herramienta monitorix) y color verde (el promedio de las tres herramientas) Fuente: Elaboración propia.

Según el promedio de los resultados de las herramientas de monitoreo, para el uso de procesador, podemos comentar de cada arquitectura lo siguiente:

1. La Arquitectura 4 es la más eficiente (41% de consumo).
2. La implementación de Docker en la arquitectura 3 logró un 10% de mejor rendimiento que la separación por capas de la arquitectura 2.
3. La arquitectura 2 tienen un resultado aceptable del 71%
4. Arquitectura 1 es la menos eficiente (94% de consumo).

RAM

La medición de memoria ram no se utilizó datos de la herramienta propia de la nube debido a que la nube nativamente no permitía capturar datos de la memoria ram, por lo tanto, se registraron mediciones solamente con la herramienta monitorix y htop.

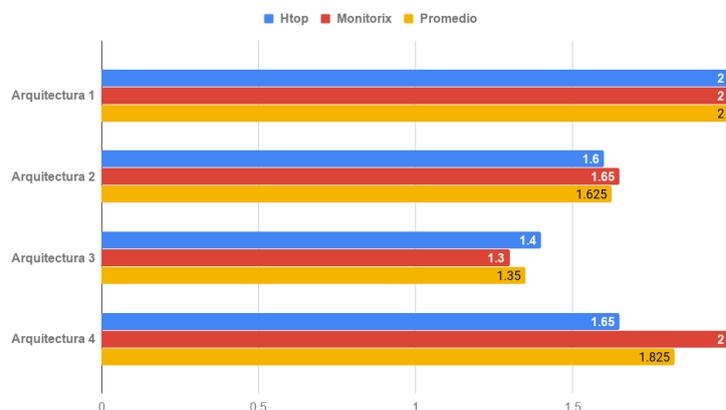


Fig. 9. Medición del uso de la memoria ram en el eje y (las arquitecturas) en el eje x (el puntaje consumido en GB, sobre un máximo de 2GB que es el mayor consumo registrado) medidos por tres herramientas, de color azul (nube), color rojo (herramienta htop) color amarillo (herramienta monitorix) y color verde (el promedio de las tres herramientas) Fuente: Elaboración propia.

1. La arquitectura que consume menos memoria ram es la arquitectura 3 consumiendo consumiendo un 33% menos que la arquitectura 1
2. La Arquitectura que más recursos consume es la arquitectura 1 logrando consumir 2Gb de memoria ram que consisten en el 50% de la memoria total disponible para el servidor (4GB)
3. La arquitectura 4 tiene una particularidad y es que la herramienta monitorix registra un uso de la misma memoria ram que la arquitectura 1, pero por el promedio se ubica en la posición 3

Almacenamiento

Debido a que el test se realizó usando el perfil visitante, este perfil es el perfil por defecto de la página no requiere inicio de sesión, no requiere guardar ningún dato por consiguiente no se hicieron consultas relevantes a la base de datos, ni se guardaron archivos y la escritura en disco duro permaneció estable.

Red

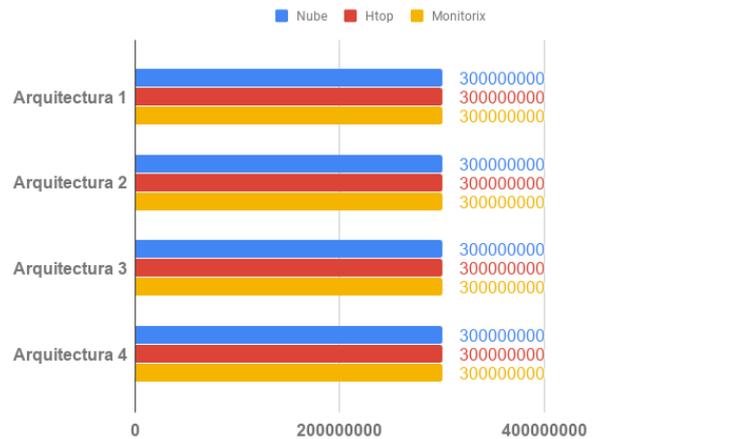


Fig. 10. Medición del uso de los paquetes de red de entrada en el eje y (las arquitecturas) en el eje x (la cantidad de paquetes de entrada expresado en bytes) medidos por tres herramientas, de color azul (nube), color rojo (herramienta htop) color amarillo (herramienta monitorix) y color verde (el promedio de las tres herramientas) Fuente: Elaboración propia.

Para las 4 arquitecturas se realizó la prueba de carga utilizando la misma cantidad de usuarios y el mismo tiempo por consiguiente la medición de los paquetes de entrada en las 3 herramientas es el mismo, se obvió poner el uso promedio debido a que por los resultados obtenidos es irrelevante ya que dio el mismo resultado 300.000.000 bytes

Prueba de Seguridad

Pregunta: ¿Qué tan seguro es el hardware y el software sobre el que se ejecuta el servicio?

Objetivo: Evaluar el software y el hardware del servidor

Para lograr el objetivo que indica evaluar el hardware al estar en nube aplica un concepto llamado seguridad compartida

Según AWS, la seguridad compartida consiste en dividir las responsabilidades en donde el cliente en la categoría IAAS se encarga de la seguridad del sistema operativo hacia arriba[41], y la empresa de nube se encarga de cubrir aspectos de red, almacenamiento, disponibilidad, funcionamiento, mantenimiento físico

Por lo tanto, por lo encontrado en la literatura nos encargaremos de testear el funcionamiento solamente del software y seguiremos el modelo de cloud.

Para el software Se realizó un test de seguridad utilizando la herramienta arachni la cual permite realizar test a aplicaciones web[42], además tiene una base de datos que incluyen las sugerencias del owasp top ten 10[43] la cual es un documento de referencia que publica las 10 principales problemas de seguridad a nivel web

El resultado de análisis de seguridad en las diferentes arquitecturas se puede apreciar en la siguiente gráfica:

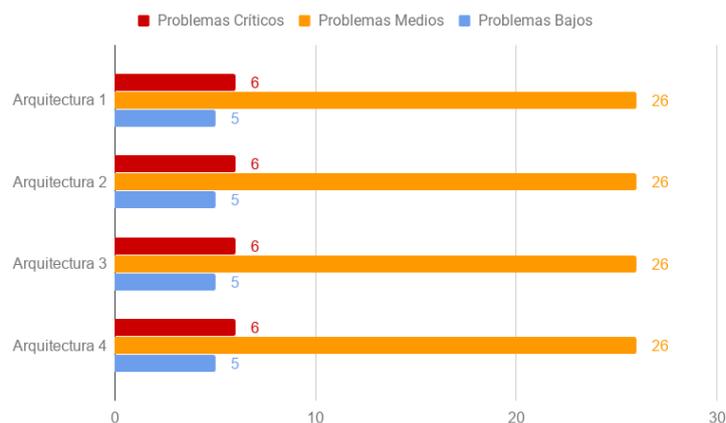


Fig. 11. Resultado del análisis de seguridad arrojado por la herramienta arachni, en el eje y (las arquitecturas) en el eje x (la cantidad de problemas de seguridad) clasificados en tres categorías, de color rojo (problemas críticos), color naranja (problemas nivel medio) color azul (problemas de nivel bajo) Fuente: Elaboración propia.

La herramienta nos permitió encontrar problemas de nivel crítico, nivel medio y nivel bajo, en todas las arquitecturas el resultado del análisis fue el mismo no se hará énfasis en todos los problemas debido a que son demasiado extenso por el contrario se realizará un resumen

Problemas de nivel crítico

Los principales problemas tuvieron relación con la autenticación dado que la información viaja sin cifrar por lo que permite el robo de identidad

Otro problema crítico consiste en la exposición de carpetas de Moodle resultado de la migración, debido a que después de realizar la migración los permisos toman por defecto el usuario apache y esto ocasiona una exposición indeseada de archivos importantes que pueden explotarse para una recolección de información y eventual ataque

Problemas de nivel medio

Dentro de los principales problemas de nivel medio encontramos problemas con relación a ataques xss, en donde estos ataques permiten la inserción de código malicioso en campos de texto que no se han filtrado[44]

Debido a que moodle maneja grandes cantidades de campos texto como lo son el de buscar cursos, buscar personas, y el login, estos campos pueden estar propensos a recibir inyecciones de código

Problemas de nivel bajo

En esta sección la herramienta realiza una búsqueda de direcciones url verificando si su respuesta es un estado 200 eso significa que la url responde al usuario positivamente, pueden o no involucrar algún problema, pero funcionan como una sugerencia a una revisión o un monitoreo que el administrador debe tener en cuenta

Económico

Pregunta: ¿Cuál es la nube más económica?

Objetivo: Lograr encontrar la nube más económica

Teniendo en cuenta el costo de una instancia en las nubes líderes mencionado anteriormente (evaluación económica de nube) para el factor económico podemos concluir lo siguiente:

1. Utilizar la arquitectura 2 y 4 tiene un costo que dobla la arquitectura 1 y 3 debido a que se requieren dos servidores para estas arquitecturas
2. El costo para hacer funcionar la arquitectura 2 y 4 es el mismo, el cambio se ve es en la tecnología que se implementa
3. El costo para la arquitectura 1 y 3 es el mismo
4. En la arquitectura 1 y 3 el crecimiento puede llegar a ser más costoso debido a que la máquina monolítica se hará cada vez más pesada
5. En la arquitectura 2 y 4 los recursos tecnológicos necesarios se compran de manera independiente lo cual puede llevar a un mejor aprovechamiento de recursos

3. Aspectos críticos y relevantes a resaltar y detallar

Transferencia de datos

Para transferir los datos de servido inhouse GIDIS a la nube de Google Cloud, se implementó la herramienta “rsync”[45].

Para seleccionar el sistema operativo del experimento, se planteó como premisa: ser soportado de manera oficial por las 3 nubes líderes, por lo que se investigó las alternativas que ofrecía cada nube, los resultados se resumen en el siguiente cuadro.

Comparativo SO disponible en las diferentes nubes

	Amazon Linux	CentOs	Container Optimized Os	Core Os	Debian	Kali Linux	Opensuse	Oracle Linux	Red Hat	SUSE	Ubuntu	Windows Server
Google Cloud	✗	✓	✓	✗	✓	✗	✗	✗	✓	✗	✓	✓
AWS	✓	✓	✗	✗	✓	✓	✗	✗	✓	✓	✓	✓
Azure	✗	✓	✗	✓	✓	✗	✓	✗	✓	✓	✓	✓

Fig. 12. Comparación de sistemas operativos soportados oficialmente por las nube. Fuente: Elaboración propia.

Para no aumentar los costos del experimento, se descartaron los SO privativos, quedando como opciones CentOS, Debian y Ubuntu. Al final se eligió CentOS como SO, debido a que posee un tiempo de soporte más prolongado (hasta 2024)[46].

Las nubes no incluyen monitoreo de memoria ram nativamente

Cada empresa de nube según la NIST (institución de estándares y tecnología) tiene la obligación de prestar el servicio de monitoreo de recursos[47] pero debido a que los servidores son virtualizados para acceder al uso de la ram se necesitan instalar agentes externos como por ejemplo stackdriver monitoring[48] o para aws tiene su propia herramienta[49], alternativamente para cada vez que se cambie de nube no deba aprender una herramienta nueva, existen herramientas que funcionan para cualquier nube como por ejemplo monitorix; Es de vital importancia escoger cuál herramienta de monitoreo se utilizará antes de salir a producción para no perder los registros históricos de uso de recursos.

4. Resultados obtenidos y su impacto

Plan de migración

Dentro de los resultados obtenidos se encuentra el plan de migración, compuesto por 8 etapas. La finalidad de cada etapa es poder cumplir unas premisas para poder continuar, además, si surgen inconvenientes, siempre se puede regresar a la etapa anterior. Las etapas son las siguientes:

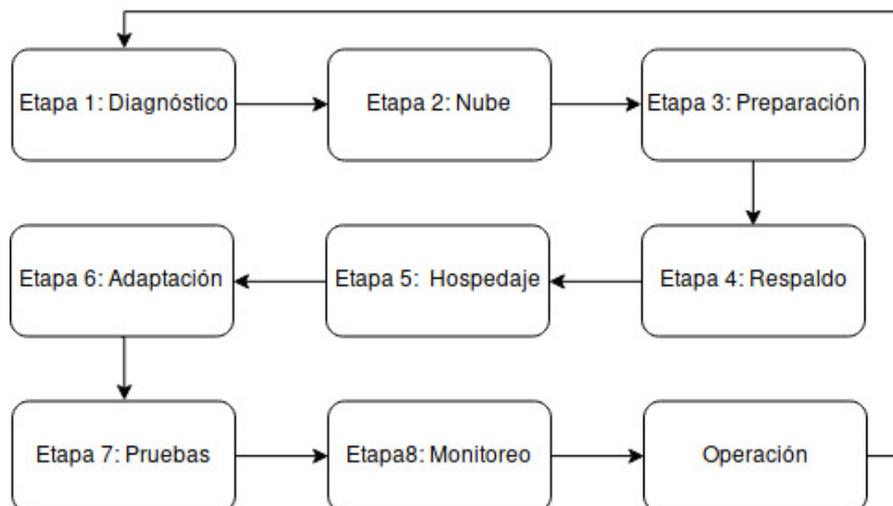


Fig. 13 Representación de las etapas del plan de migración Fuente: Elaboración propia

Etapa 1 (diagnóstico): Para superar esta etapa es necesario (i) conocer la ubicación de los archivos a migrar, (ii) tener unas estadísticas del monitoreo de los recursos actuales y (iii) conocer la versión de las dependencias, si alguna de estas premisas no se cumple, se pueden implementar herramientas de monitoreo de software libre y regresar a operación mientras se recolectan los datos.

Etapa 2 (nube): La finalidad de esta etapa es escoger el proveedor y el servicio de nube a contratar. Las estadísticas de monitoreo del diagnóstico, sirven para estimar los recursos base, y según nuestro criterio (branding, economía, rendimiento y seguridad), proceder a seleccionar la nube más adecuada a nuestra necesidad.

Etapa 3 (preparación): La finalidad de esta etapa es esperar lo necesario hasta tener una ventana de mantenimiento.

Etapa 4 (respaldo): El objetivo es realizar un respaldo de los datos (en la etapa de diagnóstico conocimos los directorios críticos).

Etapa 5 (hospedaje): Realizamos un hospedaje en nube (lift and shift)[35], donde nuestra aplicación debe poder funcionar, pero ahora sobre infraestructura cloud.

Etapa 6 (adaptación): En esta etapa debemos actualizar las dependencias que sean necesarias y realizar los cambios de arquitectura ideales para nuestra necesidad.

Etapa 7 (pruebas): Debemos testear la funcionalidad, rendimiento y seguridad de la plataforma, esta etapa se encarga de probar la calidad de la aplicación antes de salir a producción.

Etapa 8 (monitoreo): Finalmente, se activan las herramientas de monitoreo que se consideren necesarias, y se sale a operación.

Mediciones año por año

Se realizaron mediciones utilizando la herramienta Google Analytics para constatar las visitas por mes y por año, en la siguiente gráfica se aprecia un comparativo inicialmente de las visitas del 2017 contra 2018.

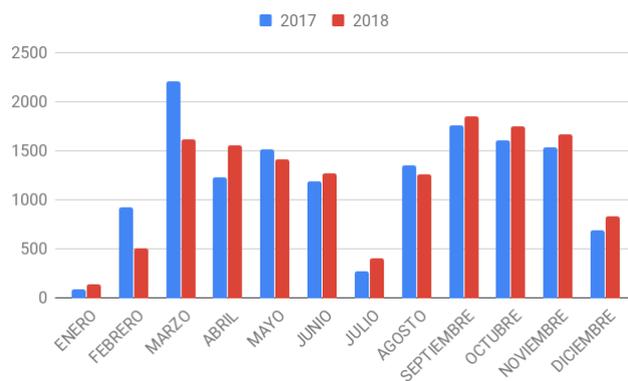


Fig. 14. Medición de las visitas de usuarios en el año 2017 vs el año 2018 , en el eje y la cantidad de visitas , en el eje x los 12 meses del año Fuente: Elaboración propia.

Luego se realizaron mediciones tomando en cuenta los meses que lleva en producción en la nube la uvirtual versus los meses en el año 2017 y 2018.

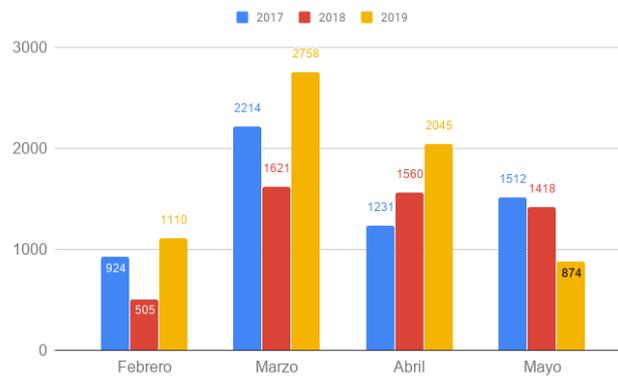


Fig. 15. Medición de las visitas de usuarios en el año 2018 vs el año 2019 con corte a mayo, en el eje y la cantidad de visitas, en el eje x los meses del año. Fuente: Elaboración propia.

Se puede constatar un alza en las visitas en el mes de marzo esto debido a que en estas fechas se realizaron los previos y muchos maestros utilizan la plataforma para evaluar con herramientas e learning como es el caso de la herramienta VPL[50] integrada actualmente, por otra parte se nota una disminución en el mes de abril debido a que se celebra semana santa y por ende es una semana en donde no se acude a clases.

5. Aprendizajes (aciertos y errores)

Inicialmente no se contaba con un monitoreo de los recursos del servidor, lo cual retrasó la migración a nube, al buscar alternativas para solucionar este inconveniente, encontramos que la comunidad de software libre brinda muchas alternativas para monitorear y administrar los recursos de una máquina, facilitándonos herramientas para continuar.

Además, los aspectos críticos anteriormente citados, se consideran aprendizajes dentro del marco del proyecto.

6. Impactos no esperados (si existieran)

Durante el desarrollo del proyecto, se generó en nuestro semillero una expectativa de los estudiantes por conocer las tecnologías de nube, esta iniciativa de ellos por aprender, nos motivó a realizar talleres y socializaciones de nuestros resultados en cloud computing, como consecuencia, se identificó que algunos estudiantes del semillero de semestres intermedios, sin haber visto materias relacionadas con programación web o base de datos, ya operaban en nube, utilizando infraestructura y servicios cloud para

realizar sus proyectos de semillero y de aula.

El objetivo del proyecto siempre giro en la órbita de migrar a cloud computing, por lo que el buen funcionamiento de la plataforma era una prueba antes de entrar en producción, pero, al migrar exitosamente a nube, se generaron ligeros cambios en la interfaz de usuario y el hallazgo de nuevas herramientas, lo que ocasionó que fuera necesario brindar una pequeña ventana de soporte y capacitación a los usuarios, la cual no se tenía inicialmente prevista en el desarrollo del proyecto.

Al finalizar el proyecto, debido al éxito de la migración, el buen rendimiento de la plataforma y los bajos costos de su implementación, la universidad decidió contratar servicios de nube, con el interés de replicar a otras aplicaciones institucionales la migración a cloud computing.

Referencias

1. Contreras MJV, Reyes JIP. Rigor y Relevancia: Un Marco Teórico y Metodológico para estudiar sistemas de información. :22.
2. Giraldo SR, Zuluaga GAO, Espinosa CL. Networking en pequeña empresa: una revisión bibliográfica utilizando la teoría de grafos. . pp. 2014;11(2):11.
3. Armbrust M, Stoica I, Zaharia M, Fox A, Griffith R, Joseph AD, Katz R, Konwinski A, Lee G, Patterson D, Rabkin A. A view of cloud computing. *Communications of the ACM*. 1 de abril de 2010;53(4):50.
4. Ercan T. Effective use of cloud computing in educational institutions. *Procedia Social and Behavioral Sciences*. 2010;5.
5. Sultan N. Cloud computing for education: A new dawn? *International Journal of Information Management*. 2010;8.
6. Kalemis K. Adapting Cloud Computing in Education: Can We Speak For an Alternative Didactic Approach in School Literacy in a Cluster Classroom for the Gifted and Talented Students? En: Uslu, F, editor. *INTERNATIONAL CONFERENCE ON EDUCATION AND SOCIAL SCIENCES (INTCESS14), VOLS I AND II*. Int Org Ctr Acad Res; 2014. p. 677-91.
7. González-Martínez JA. Cloud computing and education: A state-of-the-art survey. 2015;20.
8. Eurich M, Boutellier R. Revenue streams and value propositions of cloud-based high performance computing in higher education. En: Obaidat MS, Holzinger A, Filipe J, editores. *E-Business and Telecommunications: 11th International Joint Conference, ICETE 2014, Vienna, Austria, August 28-30, 2014, Revised Selected Papers*. Springer International Publishing; 2015. p. 61-75. (*Communications in Computer and Information Science*; vol. 554).
9. Arpacı I. Antecedents and consequences of cloud computing adoption in education to achieve knowledge management. *Computers in Human Behavior*. mayo de 2017;70:382-90.
10. Shana Z, Abulibdeh ES. Cloud Computing Issues for Higher Education: Theory of Acceptance Model. 2017;12(11):17.
11. Workineh M, Garcia NM, Midekso D. Cloud computing as technological solutions for higher education institutions: Adoption readiness assessment model: Reseach in-progress. En: 2017 International Conference on Computing Networking and Informatics (ICCNI). 2017. p. 1-7.
12. Jiang K, Song Q. A Preliminary Investigation of Container-Based Virtualization in Information Technology Education. En: *Proceedings of the 16th Annual Conference on*

Information Technology Education [Internet]. New York, NY, USA: ACM; 2015 [citado 19 de mayo de 2019]. p. 149–152. (SIGITE '15). Disponible en: <http://doi.acm.org/10.1145/2808006.2808021>

13. Colburn TR, Shute GM. Metaphor in computer science. *Journal of Applied Logic*. diciembre de 2008;6(4):526-33.

14. Calarco G. On the effectiveness of Linux containers for network virtualization. *Simulation Modelling Practice and Theory*. 2013;17.

15. Xavier MG, Neves MV, Rossi FD, Ferreto TC, Lange T, De Rose CAF. Performance Evaluation of Container-Based Virtualization for High Performance Computing Environments. En *IEEE*; 2013 [citado 3 de mayo de 2018]. p. 233-40. Disponible en: <http://ieeexplore.ieee.org/document/6498558/>

16. MERKEL D. Docker: Lightweight Linux Containers for Consistent Development and Deployment. 2014;16.

17. Xavier MG, Neves MV, De CAF. A Performance Comparison of Container-based Virtualization Systems for MapReduce Clusters. 2014;8.

18. Costache C, Machidon O, Mladin A, Sandu F, Bocu R. Software-defined networking of Linux containers. En *IEEE*; 2014 [citado 3 de mayo de 2018]. p. 1-4. Disponible en: <http://ieeexplore.ieee.org/document/6955310/>

19. Felter W, Ferreira A, Rajamony R, Rubio J. An Updated Performance Comparison of Virtual Machines and Linux Containers. :12.

20. Ruiz C, Jeanvoine E, Nussbaum L. Performance evaluation of containers for HPC. 2015;14.

21. Arnold B, Baset SA, Dettori P, Kalantar M, Mohamed II, Nadgowda SJ, Sabath M, Seelam SR, Steinder M, Spreitzer M, Youssef AS. Building the IBM Containers cloud service. *IBM Journal of Research and Development*. marzo de 2016;60(2-3):9:1-9:12.

22. Autonomous Discovery and Management in Virtual Container Clusters - OUP Journals & Magazine [Internet]. [citado 19 de mayo de 2019]. Disponible en: <https://ieeexplore.ieee.org/document/8205866>

23. Cloud Computing Trends: 2018 State of the Cloud Survey [Internet]. [citado 19 de mayo de 2019]. Disponible en: <https://www.rightscale.com/blog/cloud-industry-insights/cloud-computing-trends-2018-state-cloud-survey>

24. Federated Application Management in Docker Enterprise Edition - Docker Blog [Internet]. [citado 19 de mayo de 2019]. Disponible en: <https://blog.docker.com/2018/06/federated-application-management-in-docker-ee/>

25. Customer Success Stories | Docker [Internet]. [citado 19 de mayo de 2019]. Disponible en: <https://www.docker.com/customers>

26. Gartner Magic Quadrant for Cloud Infrastructure as a Service 2018 – BMC Blogs [Internet]. [citado 18 de febrero de 2019]. Disponible en: <https://www.bmc.com/blogs/gartner-magic-quadrant-cloud-iaas/>

27. Infraestructura global [Internet]. Amazon Web Services, Inc. [citado 19 de mayo de 2019]. Disponible en: <https://aws.amazon.com/es/about-aws/global-infrastructure/>

28. Regiones de Azure | Microsoft Azure [Internet]. [citado 19 de mayo de 2019]. Disponible en: <https://azure.microsoft.com/es-es/global-infrastructure/regions/>

29. Ubicaciones mundiales: regiones y zonas | Google Cloud [Internet]. [citado 19 de mayo de 2019]. Disponible en: <https://cloud.google.com/about/locations/#regions-tab>

30. GCP ping [Internet]. [citado 18 de febrero de 2019]. Disponible en: <http://www.gcping.com/>

31. Azure Speed Test 2.0 [Internet]. [citado 19 de mayo de 2019]. Disponible en: <https://azurespeedtest.azurewebsites.net/>

32. AWS Inter-Region Latency Monitoring [Internet]. [citado 4 de febrero de 2019]. Disponible en: <https://www.cloudping.co/>

33. Mide el rendimiento con el modelo RAIL | Web [Internet]. Google Developers. [citado

- 19 de mayo de 2019]. Disponible en: <https://developers.google.com/web/fundamentals/performance/rail?hl=es>
34. A data-centric design for n-tier architecture - ScienceDirect [Internet]. [citado 19 de mayo de 2019]. Disponible en: <https://www.sciencedirect.com/science/article/pii/S0020025502003778?via%3Dihub>
35. Migración a la nube con Amazon Web Services [Internet]. Amazon Web Services, Inc. [citado 19 de mayo de 2019]. Disponible en: <https://aws.amazon.com/es/cloud-migration/>
36. Kim T-H, Jiang K, Rajput VS. Adoption of Container-Based Virtualization in IT Education. En: 2016 ASEE Annual Conference & Exposition Proceedings [Internet]. New Orleans, Louisiana: ASEE Conferences; 2016 [citado 19 de mayo de 2019]. Disponible en: <http://peer.asee.org/26530>
37. 5 Test Measures - Testing Cloud Services [Book] [Internet]. [citado 4 de febrero de 2019]. Disponible en: <https://learning.oreilly.com/library/view/testing-cloud-services/9781492000037/ch05.html>
38. Aceto G, Botta A, de Donato W, Pescapè A. Cloud monitoring: A survey. Computer Networks. junio de 2013;57(9):2093-115.
39. htop - an interactive process viewer for Unix [Internet]. [citado 19 de mayo de 2019]. Disponible en: <https://hisham.hm/htop/>
40. Monitorix :: Documentation [Internet]. [citado 19 de mayo de 2019]. Disponible en: <https://www.monitorix.org/documentation.html>
41. Modelo de responsabilidad compartida – Amazon Web Services (AWS) [Internet]. Amazon Web Services, Inc. [citado 19 de mayo de 2019]. Disponible en: <https://aws.amazon.com/es/compliance/shared-responsibility-model/>
42. Home - Arachni - Web Application Security Scanner Framework [Internet]. [citado 19 de mayo de 2019]. Disponible en: <https://www.arachni-scanner.com/>
43. Category:OWASP Top Ten Project - OWASP [Internet]. [citado 19 de mayo de 2019]. Disponible en: https://www.owasp.org/index.php/Category:OWASP_Top_Ten_Project
44. Cross-site Scripting (XSS) - OWASP [Internet]. [citado 19 de mayo de 2019]. Disponible en: [https://www.owasp.org/index.php/Cross-site_Scripting_\(XSS\)](https://www.owasp.org/index.php/Cross-site_Scripting_(XSS))
45. rsync(1) - Linux man page [Internet]. [citado 19 de mayo de 2019]. Disponible en: <https://linux.die.net/man/1/rsync>
46. FAQ/General - CentOS Wiki [Internet]. [citado 19 de mayo de 2019]. Disponible en: <https://wiki.centos.org/FAQ/General#head-fe8a0be91ee3e7dea812e8694491e1dde5b75e6d>
47. Mell P, Grance T. The NIST Definition of Cloud Computing. :7.
48. Supervisión de nube híbrida con Stackdriver Monitoring | Stackdriver Monitoring | Google Cloud [Internet]. [citado 19 de mayo de 2019]. Disponible en: <https://cloud.google.com/monitoring/>
49. Monitorización de métricas de memoria y disco para instancias de Linux de Amazon EC2 - Amazon Elastic Compute Cloud [Internet]. [citado 19 de mayo de 2019]. Disponible en: https://docs.aws.amazon.com/es_es/AWSEC2/latest/UserGuide/mon-scripts.html
50. promedia ufps. Virtual programming Lab [Internet]. [citado 19 de mayo de 2019]. Disponible en: <https://www.youtube.com/watch?v=qfx0szRlvZA&feature=youtu.be>

Experiencias recientes en la mejora de la metodología para gestión de proyectos de infraestructura TIC en el Centro de Informática de la Universidad de Costa Rica

Fabiola Rodríguez Alfaro^a, Luis Loría Chavarría^a

^aCentro de Informática, Universidad de Costa Rica,
San Pedro de Montes de Oca, Costa Rica

fabiola.rodriguez@ucr.ac.cr luis.loria@ucr.ac.cr

Resumen. En la práctica, los proyectos de desarrollo o mejora de infraestructura TIC varían notablemente con respecto a proyectos de desarrollo de software o de otro tipo. La infraestructura TIC generalmente involucra cambios en civiles, mecánicos, eléctricos y de comunicaciones entre otros; la implementación de variaciones en estas infraestructuras generan varios retos siendo la continuidad del servicio TIC el más crítico y a este se le suman todos los retos desconocidos para involucrados en la gestión y ejecución de proyectos de este tipo. Existen algunas acciones, controles y ajustes que se pueden adoptar para mejorar la ejecución y los resultados, eliminando la incertidumbre mediante la adopción de buenas prácticas aprendidas así como la adaptación del desarrollo ágil y sus beneficios, que generan considerables impactos positivos en los proyectos de infraestructura TIC. Y es precisamente ese el objetivo de este artículo, compartir y dar a conocer los ajustes, controles y buenas prácticas que se han aprendido e implementado en la experiencia de gestión y desarrollo de la infraestructura TIC el entorno tecnológico a cargo del Centro de Informática de la Universidad de Costa Rica.

Palabras Clave: TIC, gestión de proyectos, infraestructura, buenas prácticas, PMI, Scrum

Eje temático: La conectividad y la infraestructura en la universidad digital.

1. Introducción

Este artículo plantea compartir las experiencias en gestión de proyectos de infraestructura de las Tecnologías de la Información y Comunicaciones (TIC) desde el Centro de Informática (CI), el departamento TIC, de la Universidad de Costa Rica (UCR). Es una óptica desde la experiencia de la gestión de los proyectos, pero enfocada de forma exclusiva en los proyectos de infraestructura TIC. Estos proyectos tienen comportamientos diferenciadores versus los proyectos de desarrollo de software, siendo estos la forma de agrupación más típica en los departamentos de TIC de las Universidades.

Esta diferencia conlleva que los gestores de proyectos traten de ejecutar las iniciativas en infraestructura TIC de la misma forma que desarrollos de software, lo que provoca enfrentarse a dificultades no esperadas, el nivel de incertidumbre suele ser muy alto cuando se enfrentan a las tareas específicas dentro de un proyecto, por ejemplo el reemplazo de tableros eléctricos en un centro de datos o un punto de presencia de red. Es allí donde este artículo reside, en la experiencia de estas situaciones inusitadas, poco o nulamente diseñadas y cómo se ha aprendido a sortearlas e incluso a plantear una metodología de gestión diferente que garantice el adecuado balance de las variables de un proyecto.

2. El contexto universitario y TIC

La Universidad de Costa Rica (UCR) es un centro de estudios superiores con instalaciones distribuidas en toda Costa Rica, cuenta con 8.452.462,09 m² de terrenos y 661.004,81 m²[1] de construcción dividida 12 sedes y recintos, 13 facultades, 46 Escuelas, 48 Unidades de Investigación (35 centros, 13 institutos), 2 estaciones experimentales, 19 fincas y reservas biológicas, 6 museos, 22 bibliotecas y 1 planetario, con una comunidad universitaria compuesta por 43322 estudiantes[1] y 10000 funcionarios usuarios de los servicios TIC, esta población está conformada por estudiantes, docentes, investigadores, administrativos y personas bajo algún rol de pasantías, estudiantes no formales o contratos temporales. La UCR destinó en los últimos 5 años para el presupuesto del Centro de Informática (CI) \$4 millones en promedio al año, y \$6 millones a la Comisión Institucional de Equipamiento (CIEq), esto según datos de la Dirección del CI. Estas son las dos principales fuentes de proyectos de inversión en TIC para la institución, esto sin considerar otros movimientos presupuestarios realizados por unidades académicas, Vicerrectorías o las inversiones de presupuestos autogenerados.

El CI tiene a su haber 96 personas en planta divididas en 11 áreas y unidades, a decir: Dirección, Unidad de Gestión de las Adquisiciones (UGA), Unidad de Riesgo y Seguridad (URS), Unidad de Calidad y Mejora continua (UCM), Unidad de Administrativa y de Recurso (UAR), Área de Gestión de Usuarios (AGU), Área de Gestión de Infraestructura (AGI), Área de Gestión de Comunicaciones (AGC), Área de Gestión de Servicios (AGS), Área de Captación y Promoción (ACP), Área de Desarrollo de Software (ADS) y Área de Investigación y Desarrollo (AID); los bloques se pueden ver en la figura 1.

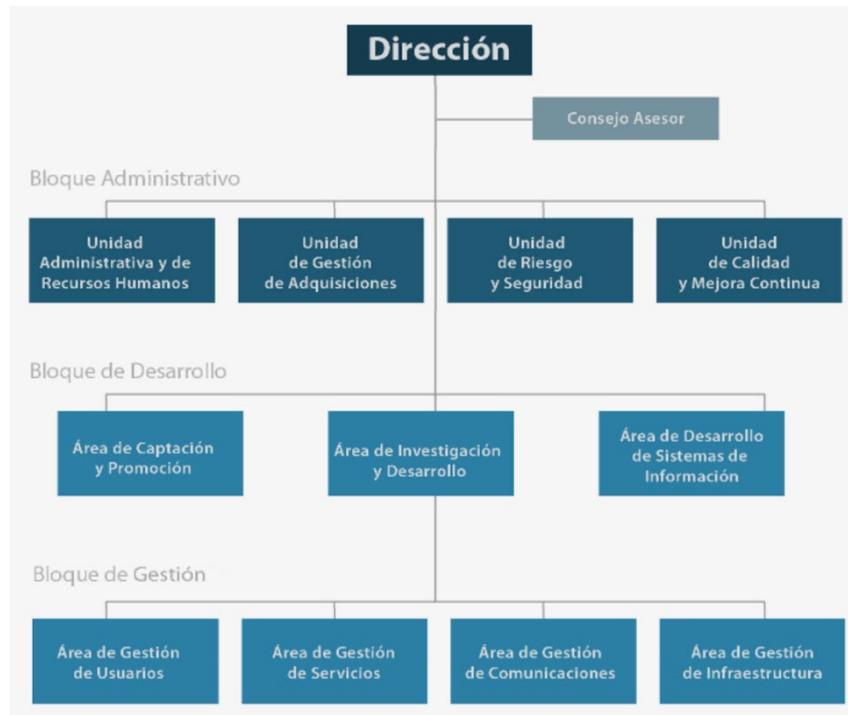


Fig. 1. Estructura organizacional del Centro de Informática de la Universidad de Costa Rica. Se puede consultar la página <https://ci.ucr.ac.cr/node/94> para conocer más sobre la organización.

Los bloques de gestión y desarrollo conforman una dupla para trabajar los desarrollos y la gestión de las plataformas, con todo el apoyo del bloque administrativo y de unidades universitarias externas al CI.

3. Metodología de gestión de proyectos utilizada

El CI posee una metodología para la gestión de proyectos basada en el PMBOK[2] del PMI. Esta herramienta de gestión utilizada en la organización, se implementó analizando los procesos y creando los instrumentos respectivos. Desde el 2017 se incursionó con la incorporación del marco de desarrollo ágil Scrum[3] para los proyectos de desarrollo de software; el cambio se notó de inmediato en las dinámicas de los proyectos de software así como en los resultados, aunado a esto, para este tipo de proyectos se rediseñaron instrumentos o herramientas y también se incorporaron dinámicas de Experiencias de Usuario, denominado como UX[4], como parte del proceso de desarrollo. Es decir, se utilizan elementos de PMI con dinámicas de Scrum y la aplicación de UX, todo un cambio en la metodología adaptada de forma específica a la naturaleza de los proyectos.

Sin embargo, para los proyectos de infraestructura se mantenía la metodología

tradicional de PMI con los instrumentos originales adaptados al CI. Así entonces, de acuerdo con la experiencia acumulada en los últimos ocho años, los proyectos de infraestructura TIC también presentan aspectos diferenciadores de frente a la metodología original existente, y requerían al igual que en el caso de los proyectos para desarrollo de software, una adaptación y mejora en la metodología y enfoque de ataque a las problemáticas.

Bajo este contexto se trabajó muy fuerte a finales del 2017 y todo el 2018 en la adaptación de PMI y Scrum versus los procesos organizacionales y las experiencias en proyectos de infraestructura TIC para generar una adaptación y mejora a la metodología existente, así como cambios y creación de instrumentos y documentación específica. Se logró.

4. Proyectos de infraestructura TIC

En la historia reciente, la industria TIC ha generado cambios acelerados en las distintas topologías y tecnologías utilizadas para brindar servicios a los usuarios. Muchos de estos cambios provocan que tengamos que realizar procesos de renovación de hardware y comunicaciones, que también conllevan cambios en la infraestructura civil y electromecánica que los sustenta. Si a esto último le agregamos la creciente dependencia de los usuarios hacia las tecnologías de información y comunicación, obtenemos la necesidad de proveer una infraestructura base robusta y actualizada, que cumpla con los requisitos de velocidad, disponibilidad y escalabilidad.

Los proyectos de infraestructura TIC incluyen por lo general labores de tipo civil, electromecánico, de comunicaciones y/o de monitoreo. Pueden abarcar desde la construcción de canalización subterránea para fibra óptica, construcción y equipamiento de cuartos de comunicación, hasta mejoras o construcciones de centros de datos. Son proyectos retadores, porque abarcan actividades de distintas índoles, pero que al ser sus productos destinados al área TIC, deben ser liderados por las oficinas encargadas de los servicios TIC.

De acuerdo con lo anterior, se requiere incorporar a estos proyectos personal con experiencia en las distintas áreas que abarcan el proyecto, para buscar una solución de calidad. Esto puede ser retador, debido a que no siempre se cuenta con personal de planta capacitado en todos los ámbitos requeridos. A pesar de lo anterior, es posible crear alianzas a lo interno de la institución, que faciliten la asesoría necesaria para ir generando experticia en el personal TIC.

En los últimos años, el Centro de Informática de la UCR, ha dedicado gran parte de tiempo y esfuerzo de un grupo de funcionarios, a renovar la plataforma de infraestructura TIC, la cual se encontraba desactualizada. Se han realizado mejoras a nivel de red externa de fibra óptica, cuartos de distribución de red, cuartos MDF y centros de datos. Partimos desde este esfuerzo reciente, para compartir algunas de las experiencias en el diseño y gestión de proyectos para infraestructura TIC, con el objetivo de exponer los hallazgos positivos y negativos que nos ha llevado la inversión de aproximadamente \$3 millones en los últimos 8 años, a través de 15 proyectos y sus respectivas adquisiciones.

5. Experiencias en la gestión de los proyectos de infraestructura TIC

Es desde esta base de números y datos que compartimos algunas experiencias recientes que podemos resumir, acerca del diseño y gestión de proyectos de infraestructura TIC, las cuales podemos exponer a través de 7 diferentes ópticas:

1. Carencias de diseño: después de experimentar problemas de avance en varios proyectos de este tipo, llegamos a la conclusión de que muchos de los problemas que se presentaron durante la ejecución de esos proyectos, se originaron en la falta de una buena planificación y diseño. Se detectó que la mayoría de los proyectos o contrataciones que forman parte de los proyectos de infraestructura tecnológica eran carentes de un diseño claro y concreto, es decir, a pesar de que se contaba con un alcance definido y objetivos claros, no se ahondaba lo suficiente en los diseños específicos según el tipo de obra, a decir: civil, electromecánica, telecomunicaciones, entre otras. Esta situación provocó ausencias significativas en la descripción de los requerimientos técnicos, los cuales crearon desbalances en las variables críticas de los proyectos, como lo son el tiempo, costo y calidad.
 - a. Caso: para el proyecto de construcción de un área encapsulada de un centro de datos, se dio el caso de que se planteó un proyecto de bajo costo y corto tiempo, que se encargaría de preparar la parte civil para dar las condiciones ambientales y estructurales para alojar una nueva plataforma de procesamiento y almacenamiento. Se planteó muy brevemente lo requerido, para cumplir con los estrechos tiempos que se tenían para la ejecución, lo cual generó el surgimiento de muchos imprevistos y retrasos durante la ejecución, ya que no se interiorizó ni documentó adecuadamente y con suficientes detalles lo que se requería. Esta problemática se repitió en algunas otras ocasiones, lo cual generó que se comenzará a buscar posibles soluciones.
 - b. Solución implementada: el diseño y planificación deben trabajarse en ciclos, donde se vaya afinando cada vez más la descripción y la logística del proyecto. Se debe ir incorporando a los involucrados necesarios, para que con cada repaso de la documentación se logre plasmar y perfeccionar la totalidad de las aristas del proyecto. Este afinamiento permite reducir los imprevistos y la materialización de riesgos del proyecto, además de que se logra que los involucrados internos tengan un dominio mental y documental sobre lo que está por venir, lo cual mejora su desempeño durante la ejecución.
2. Bajo nivel de participación de los involucrados: La naturaleza de los proyectos que nos atañe dicta involucrar varias disciplinas, es decir, la participación y colaboración heterotécnica es sustancial. Sin embargo, la baja, nula o tardía participación de los involucrados en los proyectos para infraestructura TIC tiene un nivel de impacto importante. De allí que la consideración e involucramiento activo de todos los participantes internos y externos desde el inicio se debe considerar requisito indispensable para aprobar el inicio del proyecto.
 - a. Caso: para algunos proyectos sucedió que una persona se encargaba de generar las especificaciones técnicas del proyecto individualmente

- o con sólo la colaboración de un involucrado técnico. Esto hacía que fueran muy rápidas las etapas de planificación y diseño, pero que la ejecución se viera afectada negativamente por la ausencia de conocimiento y dominio de los involucrados sobre las actividades y logística del proyecto. Además, el hecho de no involucrar a personas de todas las áreas de conocimiento, fue un desacierto, porque el personal que sí estuvo involucrado tenía nula o muy poca experiencia en la parte civil o electromecánica, lo cual generó debilidades en la ejecución y cambios para procurar cubrir las falencias en el alcance.
- b. Solución implementada: los involucrados deben definirse y ser tomados en cuenta desde la etapa inicial y durante todas las demás etapas del proyecto. Para ello debe existir una plantilla, donde se guarden los datos de cada involucrado y se establezca el rol de cada uno para el resto del proyecto. Se han logrado crear vínculos con las unidades encargadas de labores civiles y eléctricas de la Universidad, para obtener el apoyo de personal capacitado en esas áreas cuando es necesario, lo que ha fortalecido mucho las etapas de diseño y planificación y por consiguiente la ejecución y el seguimiento de los proyectos.
3. Imprevistos en los procesos constructivos y desarrollo: Toda construcción en su desarrollo presentará algún tipo de imprevisto, es lo usual, pero, ¿si es lo usual por qué no preverlo? Es exactamente lo que se trabajó, prever esas situaciones a través de labores previas.
 - a. Caso: cuando se realizan modificaciones civiles sobre edificios antiguos, donde no se cuenta con planos completos de tuberías, es muy posible que al picar paredes surjan inconvenientes a partir de la aparición o ruptura de tuberías existentes no documentadas. Si esto no se prevé de alguna forma, el hallazgo generará retrasos importantes al no contar con las medidas correctivas correspondientes.
 - b. Solución implementada: Las acciones preventivas son:
 4. Reuniones e inspecciones diarias cortas y concretas con los involucrados según el tipo de trabajo que se realiza. Las inspecciones deben documentarse, firmarse y archivarse, ya que podría ser necesario consultarlas en caso de dudas o reclamos posteriores.
 5. Una sesión de trabajo semanal con todos o la mayoría de los involucrados, donde se documente el avance y se resuelvan posibles dudas sobre lo que está por venir. Estas sesiones deben ser cortas y concretas, ya que complementan lo que se realiza durante las inspecciones.
 6. Además, los trabajos o acciones por realizar deben pasar por dos revisiones previas donde se incluye el diseño del trabajo, la simulación de las acciones para repasar la logística y la preparación de la lista de materiales y herramientas requeridos.
 7. Retrasos en los cambios y/o ajustes: Las decisiones sobre los cambios y/o ajustes no se tomaban a tiempo, era muy complejo y normalmente al tomar la decisión, el tiempo del proyecto se había perjudicado notablemente. Esta situación era propiciada por los siguientes factores:

- i. Los tomadores de decisión no contaban con toda la información del caso.
 - ii. Los cambios se plantean muy cercanos a la posible ejecución de los mismos.
 - iii. No se contaba con el tiempo de los tomadores de decisión para visitar la obra, los insumos y decidir.
 - iv. Floja determinación de opciones y sus riesgos al momento de presentar un cambio o ajuste.
8. Es entonces, mitigando estos factores que se puede atender de mejor forma un cambio o ajuste; esto con opciones mejor estructuradas que contemplen riesgos así como las acciones concretas a realizar.
9. Caso: existe tubería sobre el cielo raso que dificulta respetar la altura a la cual debe instalarse el nuevo cielo o hay tubería de agua que pasa por encima de una cuarto de distribución de red, que podría generar inundación y daño en equipos de comunicación en caso de ruptura. Se retrasan las obras mientras se comunica el problema a los encargados de área y gestores y mientras se logra coordinar una visita al sitio. La decisión no se toma de inmediato durante o después de la visita, porque no se le brinda prioridad ni seguimiento adecuado.
 - a. Solución implementada: el mayor cambio es prever con anterioridad esos eventos, para tener claro el camino a seguir y elevarlos a los tomadores de decisiones lo más pronto posible. Finalmente, se propicia la delegación de decisiones de varios niveles, con lo cual situaciones muy básicas y que no afecten el alcance, diseño y las variables del proyecto pueden ser tomadas por el nivel más básico. Cada cambio propuesto se documenta, para brindarle seguimiento y para que pase por los niveles de aprobación necesarios.
10. Supervisiones e inspecciones erráticas y de baja calidad: La poca constancia en el seguimiento diario sobre un proyecto de infraestructura TIC es catastrófico. Por naturaleza estos proyectos avanzan a un ritmo y con acciones vinculantes de forma muy diferente a los demás proyectos TIC. De allí que el estar vigilante de su desarrollo provoca un adecuado nivel de consciencia sobre las obras, los cambios y posibles ajustes provocando el avance más sólido incluso se logra impedir mala praxis u otras acciones negativas por parte los adjudicatarios hacia el proyecto.
 - a. Caso: En un proyecto de canalización de fibra óptica, el encargado de campo, personal técnico, no era constante con las supervisiones, específicamente: con las horas de trabajo, la evaluación del proceso, rutas atendidas, documentación de los cambios y avances. Esto nos daba como resultado muchos datos incompletos e inexactos. Por lo que no era fácil entender el avance del proyecto e incluso los cambios sugeridos.
 - b. Solución implementada: Esto se atendió de la siguiente manera:
 - i. Se diseñó y creó un instrumento para bitácora de campo, estas no existían. Además del instrumento, se diseñó el proceso de utilización y consolidación de los datos recopilados. Todo esto se realizó involucrando al personal técnico el cual aportó sus ideas e inquietudes al respecto,

- convirtiendo este cambio en un tema grupal y no de jefatura. El objetivo de este documento es registrar los acontecimientos de la supervisión de campo con el detalle que el proyecto requiere.
- ii. El personal técnico fue sensibilizado sobre la temática de proyectos y de contratación administrativa, esto desde la perspectiva del rol que posee el supervisor de campo. De esta forma se logró en el personal la conciencia sobre la importancia de su tarea y el impacto positivo o negativo de sus acciones.
 - iii. Se consolidan las bitácoras diarias en la reunión semanal del proyecto, de esta forma se toman como base de datos del proyecto.
11. Ausencia de hitos y controles de avance: Los proyectos de infraestructura TIC requieren, al igual que cualquier otro proyecto, la trazabilidad y seguimiento mínimo necesario. Es decir, los puntos de control sobre el desarrollo e incluso etapas previas como diseño deben ser parte de la gestión típica. En el caso de nuestra organización se define como irrefutable la revisión en al menos estos puntos:
- i. Previa contratación.
 - ii. Previa a la reunión inicial con adjudicatario
 - iii. Previa arranque del adjudicatario
 - iv. Reportes de avance cada 2 o 4 semanas.
12. Pero, además, dependiendo del proyecto se definen hitos previos al inicio de sub etapas para validar su alcance, cronograma, involucrados y los recursos asignados.
- a. Caso: Con la adecuación y construcción de cuartos de comunicación principales, denominados PoP, se presentaron problemas de seguimiento y de control entre la empresa adjudicada y la organización. Mucho del problema era la falta de orden del adjudicatario versus las actividades planeadas en cronograma, esto por cuanto la variación de un día a otro e incluso durante el mismo día era lo usual.
 - b. Solución implementada: Para procesos de contratación nuevos así como para proyectos se implementaron las revisiones ya señaladas, pero para el caso lo que se hizo fue generar reuniones de seguimiento y evaluación cada semana, así como una inspección diaria previa al arranque y al final del día. También se enfrentó a la empresa adjudicada dejando por escrito y aprobado el cronograma con el listado de recursos involucrados en dichas actividades. El control del avance y el seguimiento diario resultaron ser las mejores armas contra la actitud errática del adjudicatario. También, se colocó sobre la mesa la ley y reglamento que faculta a la organización a penalizar esos atrasos y malas prácticas. Así entonces, se complementó la ley con la mejoría en el proceso de control.
13. Ausencia de lecciones aprendidas: Se le atribuye a Napoleón Bonaparte la

frase "Aquel que no conoce su historia está condenado a repetirla", clara idea sobre el conocimiento de lo acontecido y la socialización del mismo. Esto es lo que se debe buscar en una organización con las lecciones aprendidas de todos los proyectos. Este proceso debe ser constante y con un contexto claro que autoexplique las razones de las acciones, así como de las decisiones.

- a. Caso: dentro un proyecto para la renovación, ampliación de canalización para fibra óptica, así como el tiraje de enlaces nuevos, se cometieron varias acciones equivocadas por no contar con toda la experiencia documentada de proyectos similares en años anteriores. Por ejemplo, previo a este proyecto no se instaló fibra óptica de forma masiva durante 5 años, sino que se instalaron enlaces por requerimientos y reparaciones muy puntuales; la documentación de la canalización y esos enlaces no era actualizada. Además, todos los técnicos ejecutaban el proceso físico de instalación de maneras diferentes con resultados también diferentes. No se contaba un método o proceso único para la instalación de la fibra óptica ni el uso de la canalización.
14. También, se identificaron problemas con el uso de las cámaras/arquetas ya que no se conocía cuales por nivel freático se mantienen llenas de agua, entonces en el momento de ejecutar la instalación se generaba un retraso.
- a. Solución implementada: Para minimizar el impacto sobre el desconocimiento sobre procesos, obras o infraestructuras, su uso y estado, es requisito realizar las siguientes actividades como parte de los proyectos:
 - i. En las actividades de diseño se involucra al personal a cargo de la última labor similar. Esto siempre y cuando no exista documentación verificada.
 - ii. Se trabaja en la definición de procesos de tareas y normalización de actividades repetitivas.
 - iii. La documentación y su verificación por parte de los involucrados del proyecto es un producto de todas las contrataciones.
 - iv. Se documenta y verifica las lecciones aprendidas del proyecto.
 - v. Se socializa entre los involucrados y sus áreas de trabajo las lecciones aprendidas en una única sesión de trabajo.
 - vi. La documentación administrativa del proyecto, así como los planos de índole civil, eléctrico y de datos se centraliza para contar con una biblioteca de datos e información base para el inicio de proyectos o contratos de proyectos.

6. Resultados obtenidos y su impacto

Los cambios propuestos fueron medidos como parte de un estudio académico donde se valoraba la aplicación de este cambio de metodología y la utilización de nuevos

instrumentos. Se evaluaron los atributos de alcance, tiempo, costo y satisfacción desde las variables de eficiencia y eficacia; para este análisis se realizaron encuestas a los colaboradores y gestores en los proyectos para infraestructura TIC, además se realizó un estudio sobre la documentación de 6 proyectos de infraestructura TIC y las experiencias documentales. Se utilizó la escala Likert [5] con valores entre 1 y 5, donde 5 es el mayor valor posible en sentido positivo.

Los resultados de la evaluación de los proyectos, con instrumentos y metodología versus el esquema anterior se puede ver en las figuras 2, 3, 4 y 5.

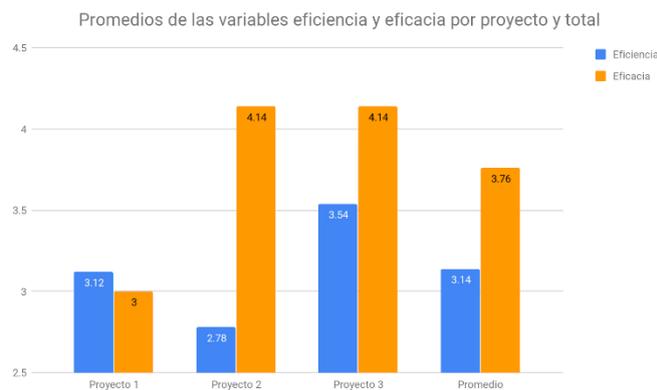


Fig. 2. Promedio de las variables eficiencia y eficacia antes de la metodología nueva. Utilizando la métrica Likert en una encuesta se mide la eficiencia y eficacia de la gestión de los proyectos de infraestructura TIC en tres proyectos. Los valores del eje X representan los proyectos y el promedio de ellos, así como en el eje Y se presenta el resultado en la escala Likert (1-5).

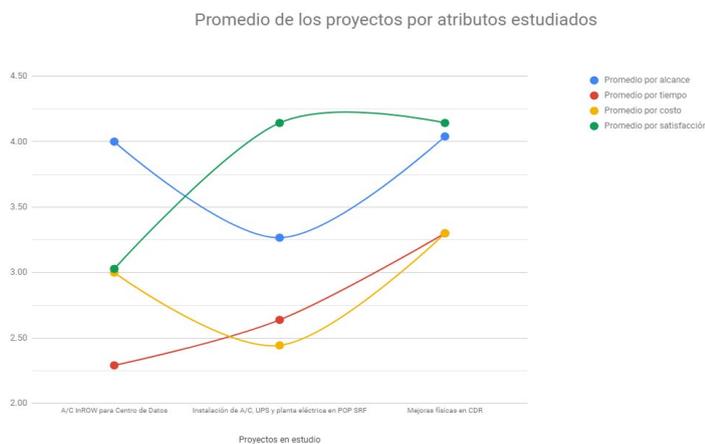


Fig. 3. Promedio de los atributos estudiados antes de la metodología nueva. Utilizando la métrica Likert en una encuesta se miden 4 atributos (alcance, tiempo, costo y satisfacción) de la gestión de los proyectos de infraestructura TIC en tres proyectos. Los valores del eje X representan los proyectos específicos y en el eje Y se presenta el resultado en la escala Likert (1-5). Se puede visualizar la comparaciones entre ellos. Todos los proyectos habían finalizado a la fecha de la encuesta a los involucrados.

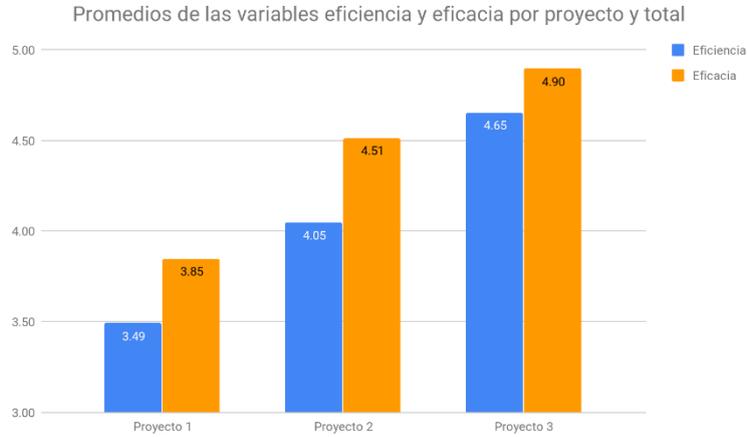


Fig. 4. Promedio de las variables eficiencia y eficacia luego de la metodología nueva. Utilizando la métrica Likert en una encuesta se mide la eficiencia y eficacia de la gestión de los proyectos de infraestructura TIC en tres proyectos en proceso muy avanzado. Los valores del eje X representan los proyectos y en el eje Y se presenta el resultado en la escala Likert (1-5). Se puede observar como los números son mejores que en la figura 2.

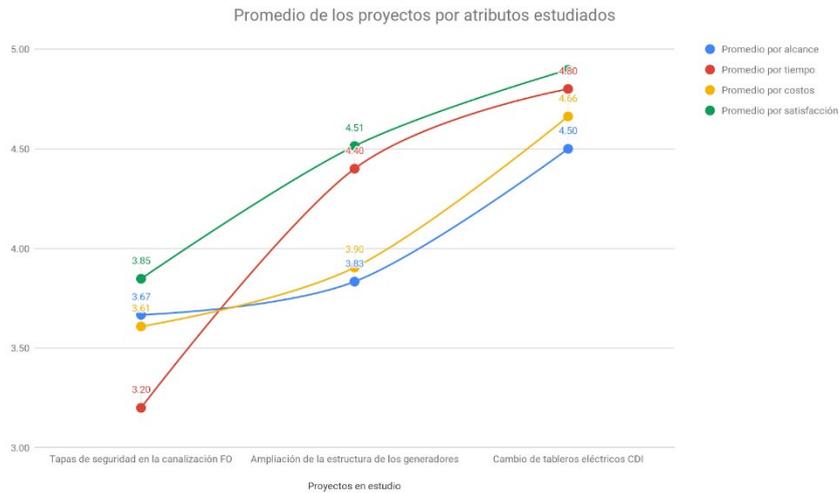


Fig. 5. Promedio de los atributos estudiados luego de la metodología nueva. Utilizando la métrica Likert en una encuesta se miden 4 atributos (alcance, tiempo, costo y satisfacción) de la gestión de los proyectos de infraestructura TIC en tres proyectos. Los valores del eje X representan los proyectos específicos y en el eje Y se presenta el resultado en la escala Likert (1-5). Se puede visualizar la comparaciones entre ellos. Estos los proyectos no habían finalizado a la fecha de la encuesta a los involucrados, sin embargo estaban muy avanzados.

Entonces, como resultados de la aplicación de la metodología nueva en tres proyectos actuales, en comparación con tres proyectos ejecutados con la metodología anterior, las variables de eficiencia y eficacia muestran cambios positivos bastante

significativos, los cuales se reflejan en la tabla 1.

Tabla 1. Comparativa entre eficiencia y eficacia al utilizar la metodología anterior y la nueva metodología

	Metodología anterior	Metodología nueva
Eficiencia	3,14	4,06
Eficacia	3,76	4,42

El impacto es importante, tomando en cuenta que los segundos proyectos estaban en los procesos de cierre durante la evaluación. Por lo tanto, queda demostrado que el resultado de los cambios incorporados en la metodología es positivo.

7. Conclusiones y recomendaciones

Con base en las experiencias vividas en el desarrollo de proyectos de infraestructura TIC, podemos entonces brindar los siguientes resultados en forma de conclusiones y recomendaciones.

7.1 Conclusiones:

1. El alcance de un proyecto de infraestructura tecnológica siempre puede ser mejor definido. Conforme el nivel de definición se ahonde, se lograrán mejores resultados durante la ejecución del proyecto, por lo que la inversión del tiempo en etapas tempranas nunca debe considerarse un desperdicio. Es necesario trabajar incrementalmente los documentos de diseño y planificación, de manera que en cada repetición se identifiquen posibles nuevos aportes o modificaciones necesarias. La discusión grupal puede traer nuevas ideas y resolver dudas en favor del dominio a profundidad de lo que está por venir.
2. Cuando se fortalece la definición del alcance del proyecto, se logran mejoras en la calidad de los resultados, ya que se incrementa la satisfacción de todas las partes involucradas, a las cuales se tuvo en cuenta a conciencia, en procura de una solución más completa y que cumpla con las expectativas generadas en las discusiones técnicas previas.
3. Con las plantillas y procesos incorporados, se logran mejoras significativas en los procesos de control y seguimiento de los proyectos, ya que las listas de chequeo, bitácoras de inspección, la gestión de cambios y los avances facilitan las labores del gestor de proyecto, en búsqueda de un mejor manejo de los costos y los tiempos planificados.
4. El diseño y planificación adecuados, así como la conciencia de lo urgente que es atender imprevistos y la conciencia sobre los posibles riesgos, generan

mejoras en los tiempos de avance y finalización de este tipo de proyectos, ya que se logra minimizar los períodos de detención de las obras.

5. El involucramiento temprano de los interesados e involucrados trae consigo mejoras en la definición del alcance y la logística, los cuales son claves en los proyectos de infraestructura tecnológica donde se deben abarcar varias áreas de conocimiento. Las discusiones y recreaciones generadas durante las reuniones de diseño entre las diferentes partes fortalecen el vínculo entre los involucrados y les brindan claridad sobre las actividades del proyecto, lo cual fortalece las labores de supervisión, inspección, seguimiento y detección de problemas o debilidades. Además, el establecimiento de roles claros permite una mejor organización del equipo de trabajo.

7.2 Recomendaciones:

1. Es recomendable mezclar metodologías estructuradas como PMI con metodologías ágiles, como SCRUM, para que existan etapas con desarrollo incremental, en el que se den pequeñas entregas o avances que fortalezcan tanto las primeras etapas, como la ejecución y seguimiento.
2. Para los proyectos de infraestructura TIC es necesario adaptar instrumentos de documentación y seguimiento, con el fin de que cumplan con la función de simplificar el trabajo de gestión y apoyar durante la etapa de ejecución. Las listas de chequeo y la gestión de cambios son fundamentales.
3. Definir procesos más claros relacionados con los proyectos y la ley de contratación, así como alinear las etapas de cada uno de los marcos de trabajo generan un mejor entendimiento del escenario desde donde se gestiona proyectos.
4. Evaluación y cambio continuo de los procesos debe ser una tarea diaria, pero este accionar debe estar interiorizado desde las jefaturas hacia los colaboradores. La adaptación y cambio debe ser una constante y no debe impactar negativamente a los grupos de trabajo. La horizontalidad con adecuada comunicación facilita estos procesos que son creativos.
5. Socializar las lecciones aprendidas tal como se mencionó anteriormente es un arma vital para la organización. Si bien es cierto desde el PMI se recomienda la utilización de instrumentos para registrar las lecciones aprendidas, esta práctica no siempre se realiza, además, es usual que la información queda allí, sólo documentada. Así entonces, realizar procesos de socialización grupal para discusión de las lecciones aprendidas es una práctica altamente recomendada.
6. Incentivar el trabajo en equipo horizontal representa una ventaja en las sesiones de trabajo creativo donde se analice y realice el diseño de proyectos así como en actividades de control y seguimiento sin olvidar el espacio para socializar las lecciones aprendidas. Se sabe que es un cambio cultural, pero es necesario para implementar cambios basados en metodologías ágiles.

8. Aprendizajes: aciertos y errores

Como se mostró en las conclusiones, recomendaciones e impacto, la experiencia es mayormente positiva, sin embargo, ese resultado se obtiene luego de malas experiencias y errores en los proyectos de infraestructura TIC. Desde los mismos y con la sinergia en el trabajo en equipo de forma horizontal, se logran procesos creativos y de discusiones propositivas.

La utilización de listas de chequeo como instrumento de control previo a tareas, así como para el diseño, representa un acierto práctico que incide en la calidad de los resultados. Este tipo de instrumentos realmente mejora la cultura de los grupos de trabajo impactando en los atributos de tiempo, costos y calidad.

El mayor aprendizaje que se puede señalar es la necesaria y adecuada participación de todas las personas involucradas en un proyecto de este tipo, donde se procuren canales de comunicación asertiva.

Agradecimientos

Los autores desean expresar su agradecimiento al equipo de trabajo que ha sido partícipe de las mejoras metodológicas aquí planteadas.

Referencias

1. UCR en cifras, <https://www.ucr.ac.cr/acerca-u/ucr-en-cifras.html>
2. Project Management Institute: Guía de los fundamentos para la dirección de proyectos (Guía del PMBOK). Project Management Institute, Pennsylvania (2017).
3. Roche, J: ¿Qué es Scrum? El método ágil para gestión de proyectos más usado en el mundo. Deloitte, <https://www2.deloitte.com/es/es/pages/technology/articles/que-es-scrum.html>
4. Aceves, LC: Metodología centrada en la Experiencia del Usuario, <http://www.luiscarlosaceves.com/udem/dweb/theelementsofuserexperience.pdf>
5. Sampieri, R: Metodología de la investigación, pp. 238 a 249. Sexta Edición. McGraw-Hill Interamericana (2014).

Glosario

1. Escala de Likert: escala psicométrica utilizada comúnmente en cuestionarios, que consta de cinco posibles valores de respuesta. Creada por Rensis Likert.
2. PMBOK (Project Management Body Of Knowledge): Guía de los fundamentos para la dirección de proyectos, es un libro que agrupa estándares, normas y buena prácticas para la gestión de proyectos, creada por el PMI.

3. PMI (Project Management Institute): Instituto de Administración de Proyectos, es una organización internacional fundada en Estados Unidos que agrupa a profesionales relacionados con la administración de proyectos. Brinda certificaciones que son reconocidas a nivel mundial.
4. Scrum: es un marco de trabajo que busca reducir la complejidad del desarrollo de productos y satisfacer las necesidades del cliente. Se basa en tecnologías ágiles de gestión y entrega productos de forma incremental.

Capítulo 3

La gestión y gobierno de las TIC con miras a la transformación digital universitaria

Trascendencia de la gestión de los servicios de TI en la vida universitaria

Graciela Cecilia Flores Reyes, Tomás Rodríguez Elizondo

Universidad Autónoma de Nuevo León, Dirección de Tecnologías de Información, Ciudad Universitaria, CP 66455, San Nicolás de los Garza, Nuevo León, México.

Resumen. La gestión de los servicios de tecnologías de información ha tomado mayor auge en los últimos años debido a que actualmente el ser humano se ve involucrado diariamente con alguna actividad relacionada a la tecnología, desde una transacción bancaria, hasta el envío de un mensaje a través de una aplicación móvil. En la comunidad universitaria, también se presenta esta tendencia, los estudiantes cursan clases en línea, interactúan en foros de discusión, entregan sus tareas de manera virtual, presentan evaluaciones en línea y consultan sus calificaciones a través de plataformas educativas, sin embargo, si estas herramientas que fueron creadas para apoyar el crecimiento educativo y la agilización de los procesos administrativos, no se gestionan de manera adecuada, se corre el riesgo de que no se proporcione la disponibilidad requerida, la capacidad adecuada o la continuidad de los servicios educativos, esto entre otros factores que ayudan a determinar que realmente se genere valor en los servicios proporcionados a la vida universitaria.

Palabras Clave: Gestión de Servicios de Tecnologías de Información.

Eje temático: Transformando los servicios hacia la mejora de la experiencia del usuario (comunidad universitaria).

1. Introducción

Los Servicios de Tecnologías de Información han estado cada vez más involucrados en la comunidad universitaria mediante sus procesos académicos, administrativos, de investigación, extensión y vinculación, se han creado con el fin de responder a las necesidades de la Universidad Autónoma de Nuevo Leon (UANL), tales como el crecimiento, la optimización de recursos, la innovación, entre otras.

Los servicios de TI, juegan hoy en día un papel muy importante en la vida del ser humano ya que diariamente interactuamos con la tecnología en un sinfín de actividades que forman parte de nuestra vida, la academia a la par de esta evolución generacional ha ido transformando la manera en que se brinda la educación desde ofrecer clases en línea hasta proporcionar acceso a un mundo de información con tan solo pulsar un botón.

Bajo estas premisas la Dirección de Tecnologías de Información (DTI) identificó lo crítico que era el área de informática para todas las dependencias Académicas y Administrativas de la UANL y se enfrentaba a los siguientes retos; con el recurso humano disponible tenía que seguir brindando servicios innovadores y de calidad, la operación actual no permitía medir el desempeño y la eficiencia de los servicios, lo cual imposibilitaba la toma de decisiones al no poder determinar oportunamente el desempeño y los costos operativos. Adicionalmente se presentaba triangulación de información, duplicidad de roles, falta de comunicación durante la atención de las solicitudes y finalmente no se tenían acuerdos establecidos con los clientes con respecto a los tiempos de entrega.

Buscando una solución a estas problemáticas, la DTI refuerza su Visión y se compromete a proporcionar Servicios de Tecnologías de Información que cumplan con las mejores prácticas y certificaciones reconocidas internacionalmente, con el fin de dar a la comunidad universitaria servicios de calidad se vio la necesidad de apoyar nuestro sistema de calidad con la implementación de marcos de referencia.

La DTI realizó un análisis sobre la gestión actual de los servicios de TI en la Institución, y cuáles eran las mejores prácticas que aportarían valor a la entrega de servicios determinando que la mejor forma de robustecer su entrega, era con la implementación de ITIL (*Information Technology Infrastructure Library por sus siglas en inglés*), el cual es un conjunto mundialmente reconocido de mejores prácticas para la Gestión de Servicios de Tecnologías de Información (ITSM), y apoyó en la identificación y mitigación de riesgos operativos, reforzando la continuidad del servicio mediante planes de contingencia, implementando la medición de la demanda, la capacidad y la disponibilidad y guiando al establecimiento de un compromiso formal con los clientes con respecto a los tiempos de solución de cada uno de nuestros servicios.

2. El cambio de gestión de servicios en la DTI

La DTI cuenta con la certificación ISO 9001 desde el año 2005, lo cual ayudó a minimizar el reto de documentar los procesos y la medición de los mismos, sin embargo, fue uno de los principales paradigmas que se encontraron en un inicio de la

implementación de ITIL e ISO20000.

Con la certificación de ISO 9001 se había establecido un Proceso de Servicios que establecía responsabilidades definidas, la creación de un help desk y el inicio de algunos indicadores aislados, sin embargo, aunque se hacía lo que estaba documentado, nadie nos aseguraba que las cosas se estuvieran realizando de la mejor manera posible.

2.1 ITIL

Considerando que ITIL, se presenta como un conjunto de buenas prácticas, es decir, un enfoque o método que ha demostrado su validez en la práctica [1] y que establece un ciclo de vida del servicio que considera las fases de: estrategia, diseño, operación, transición y mejora continua de servicios, se tomó como referencia para reestructurar el Proceso de Servicios existente.



Fig. 1. Fases del ciclo de vida de ITIL (fuente: DTI).

A mediados del 2008 cuando la DTI decidió dar el primer paso para la transformación de la estructura que se tenía para la gestión de los servicios, analizó este ciclo de vida, comenzando por la definición de políticas y objetivos, es decir el compromiso de la Alta Dirección, posteriormente, se identificó al personal involucrado con cada una de las áreas de proceso para establecer, dar seguimiento e implementar los subprocesos que eran requeridos por la mejor práctica.

Como primera fase de implementación se definieron los siguientes procesos:

Catálogo de Servicios.- Diseñado para proveer información acerca de los servicios de TI por la Dirección de Tecnologías de Información de la UANL.

Administración de Incidentes.- Ser el único punto de contacto para atender y solucionar las solicitudes relacionadas con los servicios de TI, que hacen los usuarios

de la UANL a través del CIATI buscando la satisfacción de sus actividades cotidianas.

Administración de Problemas.- Resolver la causa raíz de incidentes que son causados por errores en la infraestructura y prevenir la recurrencia de los mismos, además de minimizar el impacto adverso de los incidentes que no puedan ser prevenidos.

Administración de Cambios.- Asegurar el empleo de métodos y procedimientos estándar para manejar eficiente y oportunamente todos los cambios de tal manera que se minimice el impacto en la calidad de los servicios, así como mejorar la operación diaria de la infraestructura.

Administración de Liberaciones.- Desplegar las liberaciones al ambiente de producción y establecer el uso efectivo de Servicios de TI para brindar valor a los clientes y habilitar la transferencia a operaciones.

Administración de Configuraciones.- Tener todos los componentes de servicios e infraestructura de TI bajo el control de la Administración de Configuraciones, y por consecuencia, facilitar la administración de cambios, el manejo de incidentes y problemas, así como contribuir a la entrega y soporte de servicios de TI de calidad, permitiendo enfrentar cambios del negocio y requerimientos de los Clientes.

Administración de Niveles de Servicios.- Subproceso encargado de establecer niveles de servicios que mejoran contantemente a través de acuerdos entre la Dirección de Tecnologías de Información y las Dependencias de la UANL.

Derivado de la implementación de estos procesos se creó un punto único de contacto para atender las solicitudes y fallas en los servicios de TI el cual fue llamado "Centro Institucional en Atención de Tecnologías de Información (CIATI)", con el fin de integrar en un solo sitio los requerimientos expresados por el usuario, la documentación acerca de la atención proporcionada por el responsable del área de servicio, el seguimiento a la atención de cada reporte de servicio, así como el cierre de cada una de las solicitudes. Para lograr este cambio y poder brindar seguimiento a todas las actividades que se desprenden para la administración del mismo, se necesitó habilitar un Equipo de Operaciones de TI, además de proporcionar capacitación al personal involucrado y certificarlo en Fundamentos de ITIL; en paralelo realizar la incorporación de una mesa de ayuda y afinar el Catálogo de Servicios. Poco a poco se comenzó a identificar a la DTI como parte fundamental de la operación diaria de la universidad, por la calidad y la entrega de valor que ofrecía en sus servicios de TI.

La DTI cuenta actualmente con un Catálogo de Servicios, integrando un total de 61 servicios, los cuales se muestran en la Fig. 2:

CLASIFICACIÓN DE CATALOGO DE SERVICIOS

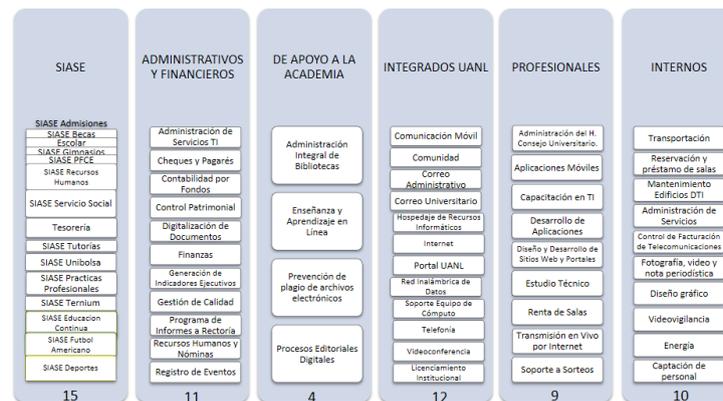


Fig. 2. Catálogo de los servicios ofrecidos por la Dirección de Tecnologías de Información. (Fuente: DTI).

Con la implementación de este primer grupo de procesos, el establecimiento de la mesa de ayuda “CIATI” y con la integración de estos subprocesos al Sistema de Calidad ya establecido en la DTI, concluyó exitosamente la primera fase del proyecto a mediados del año 2010, dando pauta a la creación de un Sistema Integral donde se conjugarían las diferentes certificaciones con las que se comenzaba a trabajar en la organización.

A partir de 2011 dio inicio a una segunda fase ya que era necesario continuar con la implementación del resto de los procesos que eran requeridos por ITIL, concentrando como segundo bloque los siguientes:

Administración de Eventos (Proceso de Operaciones).- Identificar y dar prioridad a la infraestructura, servicios, procesos de negocio y eventos de seguridad, estableciendo la respuesta adecuada a los acontecimientos tecnológicos, respondiendo a las condiciones que podrían conducir a fallas potenciales que puedan dar como resultado un incidente o problema.

Administración de la Disponibilidad.- Todos los Servicios y componentes de TI que cumplan con los requisitos de disponibilidad de los clientes de la DGI basado en la operación de la Institución y en los tiempos comprometidos de recuperación.

Administración del Portafolio de Servicios de TI.- Subproceso encargado de Gestionar los servicios derivados de los proyectos en desarrollo, catálogos de servicios y servicios retirados.

Administración de Costos de los Servicios.- Proveer una relación costo-efectividad de los recursos y activos de TI que son utilizados para proveer los servicios institucionales de Ti. E incluir el presupuesto de las acciones definidas por el Subproceso.

Mejora Continua.- Registrar, Implementar y difundir las mejoras y lecciones aprendidas de los Procesos, Proyectos y Servicios de TI, basados en las necesidades de mejora.

2.2 ISO 20000

Una vez madurados estos procesos la Alta Dirección tomó la decisión de implementar ISO/IEC 20000-1:2011, la cual es la norma internacional sobre Gestión de Servicios de TI (ITSM), publicada por ISO (Organización Internacional de Normalización) e ICE (Comisión Electrotécnica Internacional). Esta norma nos ayudó a validar y confirmar la correcta implementación de las prácticas de ITIL a través una casa certificadora.

El estándar ISO/IEC 20000 considera la implementación de las siguientes cláusulas:

- Cláusula 4. Requisitos generales del sistema de gestión de servicios
- Cláusula 5. Diseño y transición de servicios nuevos o modificados
- Cláusula 6. Procesos de entrega de servicios
- Cláusula 7. Procesos de relación
- Cláusula 8. Proceso de resolución
- Cláusula 9. Procesos de control
- Cláusula 10. Mejora continua

Como parte de la implementación de ISO/IEC 20000, fue necesario integrar algunos procesos adicionales al Proceso de Servicios para poder obtener esta certificación, tales como:

Administración de la Capacidad.- Todos los Servicios y Componentes de TI que cumplan o excedan las necesidades de capacidad requerida y acordadas con los clientes de la DTI basado en la operación de la Institución y en los tiempos comprometidos de recuperación.

Administración de la Demanda.- Entender el patrón de comportamiento de la Dirección de Tecnologías de Información y relacionar estos patrones con el impacto de la provisión de servicios de TI. Sincronizar el consumo (demanda) con la capacidad (oferta) de los recursos de TI.

Administración de la Continuidad.- Soportar los Planes de Continuidad del Negocio asegurando que los servicios de TI y capacidades técnicas puedan ser recuperadas dentro de las escalas de tiempo comprometidas.

Administración de Seguridad de la Información.- Garantizar que la implementación y operación de mecanismos de seguridad de TI está alineada con los requerimientos de seguridad de la institución a través de una administración eficiente y efectiva de los procesos de TI enfocada a la mitigación de riesgos.

Administración del Conocimiento.- El objetivo es garantizar que la información correcta es entregada en el lugar y momento adecuado a la persona indicada con el fin de que esta pueda tomar decisiones basada en información oportuna.

Relaciones con el Negocio.- encargado de establecer y mantener una buena relación de información de los servicios entre los representantes de Área de Servicio (RAS) y el cliente interno/ externo, basándose en el entendimiento del cliente/ usuario.

Administración de Proveedores.- Administrar los servicios que entregan los

proveedores, asegurando y garantizando su servicio.

Informes de Servicios Generar en plazo los informes acordados, fiables y preciso, para informar de la toma de decisiones y para una comunicación eficaz.

Una vez establecidos los procesos requeridos para la certificación de ISO 20001 era necesario determinar la estrategia de certificación, siendo elegidos como piloto de certificación dos servicios, que estratégicamente apoyaban a la comunidad universitaria, tanto a la academia y a la administración de la UANL, los cuales fueron: Administración Integral de Bibliotecas y Correo Administrativo.

Posteriormente, se fueron sumando los servicios a certificar en la Norma ISO/IEC 20000-1:2011 desde el 2014 al 2019. Y desde el pasado mes de abril la DTI cuenta con la firma de los SLA's de 27 de sus servicios certificados en la Norma ISO 20000, se listan a continuación:

- **Bloque 1**
 - Correo Administrativo
 - Administración Integral de Bibliotecas
- **Bloque 2**
 - Telefonía
 - Portal UANL
 - Enseñanza y Aprendizaje en Línea
 - Recursos Humanos y Nóminas
 - SIASE (Sistema Integral para la Administración de Servicios Educativos) Recursos Humanos
 - Gestión de Calidad
 - Comunidad
- **Bloque 3**
 - SIASE Tutorías
 - SIASE Prácticas Profesionales
 - Digitalización de Documentos
 - Internet
 - Videoconferencia
 - Generación de Indicadores Ejecutivos
 - Soporte a Sorteos
 - Administración del H Consejo Universitario
 - SIASE Unibolsa
 - SIASE Servicio Social
- **Bloque 4**
 - SIASE Educación Continua
 - SIASE Gimnasios
 - SIASE PFCE
 - SIASE Futbol Americano
 - SIASE Deportes
 - SIASE Becas
 - Red Inalámbrica de Datos
 - Transmisión en vivo por Internet

Uno de los puntos destacables de esta certificación es que se cuenta con Acuerdos de Nivel de Servicio, SLA (Service Level Agreement) firmados con el cliente, esto como parte de los objetivos que se tenían de la implementación. Algunos puntos

relevantes de este documento son:

- Alcance del Servicio
- Requerimiento de los usuarios
- Limitantes de las Funcionalidades y Características del Servicio.
- Supuestos
- Desempeño de Niveles de Servicio
- Niveles de Servicio Específicos
- Responsabilidades de la DTI
- Responsabilidades del Cliente
- Horario de Cobertura, Tiempos de Respuesta y Escalación
- Respuesta a Incidentes
- Prioridad
- Escalación
- Excepciones del Servicio para cobertura

3. La trascendencia del cambio

Para llevar a cabo una administración eficaz de la administración de los servicios y se llevase un control de los incidentes, problemas, cambios, liberaciones, configuraciones, niveles de servicio y catálogo se analizaron herramientas en el mercado que se adaptarán a la necesidades de la DTI para contar con el registro los datos, cantidad de reportes de fallas y requerimientos, nivel de impacto de fallas, dependencias atendidas, servicio más demandado, tendencias de acuerdos de servicios, medio de llegada, medio de solución, problemas registrados, cambios efectuados, cambios de emergencia y los SLA de los servicios.

Como resultado de este análisis se adquirió una herramienta para facilitar el registro de incidentes, el manejo de problemas, y su relación entre ellos y llevar así mismo, el control de los cambios de TI solicitados, las configuraciones de los elementos de configuración a través de la CMBD (Base de datos de la gestión de configuración), las liberaciones realizadas, así como la relación entre cada uno de ellos.

Adicional a esta herramienta de administración de servicios, la DTI ha generado sus propios desarrollos tecnológicos que ayudan a la gestión de los servicios que brinda a la comunidad universitaria, los cuales facilitan la entrega de valor, para proporcionar los indicadores a la Administración de Servicios, se creó el Sistema de Indicadores Ejecutivos (SIE), el cual consiste en la consulta, análisis y reporte de la información y de los datos que nos proporciona la herramienta utilizada en el CIATI, así como, conocer la cantidad de servicios ofrecidos, dependencias atendidas, nivel de impacto en las fallas, medio de llegada, medio de solución, tendencias de acuerdos de servicios, entre otros.

Los Representantes de Área de Servicio (RAS) quienes reciben el requerimiento y lo atienden hasta su cierre, documentan la asignación en sus áreas por medio de la

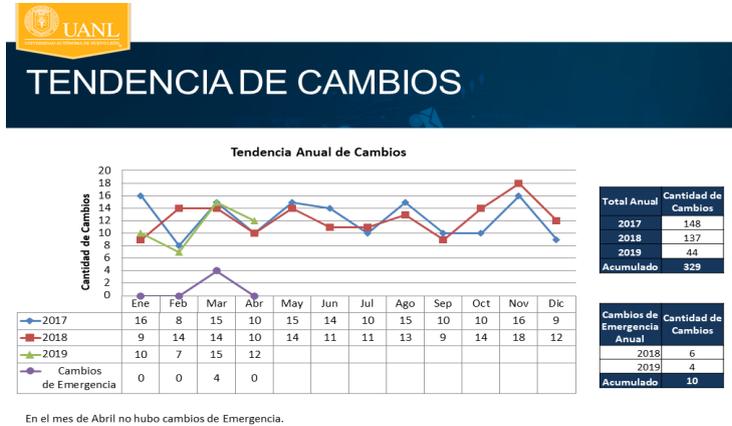
página de CIATI RAS (<https://ciati.uanl.mx>), en el cual se refleja la atención, el seguimiento y la solución de las asignaciones.

4. Principales retos y beneficios

Los principales retos que se presentaron durante este cambio de gestión de los servicios de la DTI fueron: el entendimiento de los usuarios para el registro de solicitudes por medio del CIATI, a través de una llamada telefónica, un correo electrónico u oficio y que lo identificarán como punto único de contacto, así como la aceptación de la cultura de documentar los servicios por parte del personal interno involucrado.

Como parte de los principales procesos que se desarrollaron y han generado valor son: la *administración de cambios*, ya que a través de ello se mantiene la comunicación internamente y hacia la comunidad universitaria a través de juntas de aprobación de cambios; informando por medio de correo institucional lo que la DTI realiza en sus cambios desde una actualización de parches, cambio de versión de aplicación, hasta un cambio de infraestructura central; la *administración de la capacidad* porque debido a ella y los pronósticos de la demanda de los servicios se robustece la infraestructura para garantizar el buen funcionamiento de los mismos; la *administración de la disponibilidad* que es la habilidad de un servicio o componente de TI para realizar la función requerida en un instante o periodo de tiempo determinado, la *administración de la continuidad* para soportar los Planes de Continuidad establecidos en la DTI, asegurando que los Servicios de TI y capacidades técnicas puedan ser recuperados dentro de las escalas de tiempo comprometidas y seguir garantizando el servicio cuando se presente algún acontecimiento no programado y que afecte la prestación de los servicios a la comunidad universitaria y la *Administración de la seguridad de la información* que nos apoya en gestionar la seguridad de la información de manera eficaz para todas las actividades del servicio prestado.

El comportamiento de los procesos más relevantes obtenidos se muestran a continuación:



Servicios - Estadísticas Abril 2019

Fig. 3. La tendencia de los cambios que se realizan en la DTI por medio del proceso de Administración de Cambios, ha tenido una aceptación de gran relevancia debido a que queda registrado la afectación del servicio, CI afectado, plan de implementación con horarios y responsables de la actividad, así como en determinado momento si no se cumple con las actividades, retornar al estado inicial y una autorización por parte del área de seguridad para asegurar que no cuente con vulnerabilidades en el cambio.

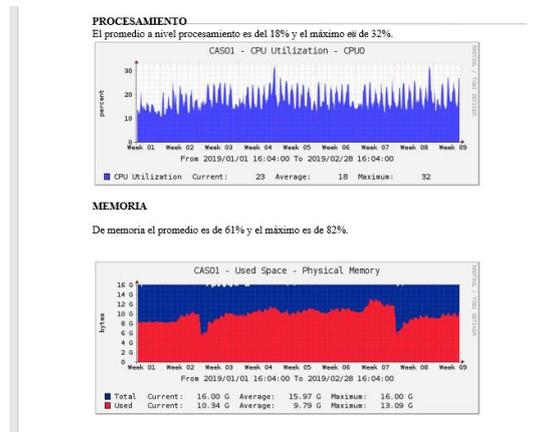


Fig. 4. La gráfica muestra la tendencia de la capacidad de procesamiento y memoria de los servidores donde se encuentra alojada la información de los servicios y tienen como objetivo concentrar los requerimientos particulares de capacidad actual y futura para los servicios que ofrece la DTI; establecer los requerimientos de monitoreo de la capacidad de los componentes involucrados en la entrega de los servicios de la DTI y establecer las estrategias para resolver los requerimientos y riesgos asociados con la capacidad, que sean financiera y legalmente viables.

		Tablero de Reporte de Disponibilidad					
		Año 2018					
		Julio	Agosto	Sep	Oct	Nov	Dic
Correo Administrativo	Disponibilidad	100.00% 98.5%	100.00% 98.5%	100.00% 98.5%	100.00% 98.5%	99.8762% 98.5%	100.00% 98.5%
	MTBF	SIN FALLAS 734 hrs	SIN FALLAS 734 hrs	SIN FALLAS 710 hrs	SIN FALLAS 734 hrs	SIN FALLAS 710 hrs	SIN FALLAS 734 hrs
	MTRS	SIN FALLAS 10 hrs	SIN FALLAS 10 hrs	SIN FALLAS 10 hrs	SIN FALLAS 10 hrs	SIN FALLAS 10 hrs	SIN FALLAS 10 hrs
Administración Integral de Bibliotecas	Disponibilidad	100.00% 98.5%	100.00% 98.5%	100.00% 98.5%	100.00% 98.5%	99.8857% 98.5%	100.00% 98.5%
	MTBF	SIN FALLAS 734 hrs	SIN FALLAS 734 hrs	SIN FALLAS 710 hrs	SIN FALLAS 734 hrs	SIN FALLAS 710 hrs	SIN FALLAS 734 hrs
	MTRS	SIN FALLAS 10 hrs	SIN FALLAS 10 hrs	SIN FALLAS 10 hrs	SIN FALLAS 10 hrs	SIN FALLAS 10 hrs	SIN FALLAS 10 hrs

Fig. 5. El tablero de la tendencia de la disponibilidad del servicio nos permite asegurar que los umbrales establecidos se pueden medir; Incluir derechos de acceso, tiempos de respuesta y disponibilidad de componentes de Servicio. Se monitorea cada componente CI registrado en la Base de Datos de Componentes, cada responsable operativo de los mismos registra en su bitácora diaria la operación de los equipos, lo cual plasma en las gráficas de disponibilidad los componentes involucrados en el Tablero de Disponibilidad. Así mismo, se toma en consideración el Análisis de Riesgos, para contemplar el nivel de impacto en la disponibilidad del servicio.

Así mismo, se tiene implementado un *Plan de Continuidad* que tiene como objetivo el de proveer instrucciones detalladas a los miembros del equipo de recuperación quienes han sido asignados a recuperar los servicios en caso de presentarse una contingencia. Se cuenta con 13 escenarios en este Plan que incluyen: fallos en diferentes ruteadores, UPS, planta de emergencia, climas de site y servers. Anualmente se realizan simulacros para probar que estos planes sean efectivos.

Además, la implementación de estos procesos nos dio como resultado los siguientes beneficios:

- Obtener una ventaja competitiva con respecto a organizaciones educativas del mismo rubro.
- Regir a la DTI bajo un modelo de calidad para gestionar la entrega y soporte de sus Servicios de TI.
- Mejorar la calidad, consistencia y reputación de los Servicios de TI en la UANL.
- Alinear los servicios de TI con la estrategia de la UANL.
- Entregar servicios de TI alineados a mejores prácticas de clase mundial.
- Definir claramente las responsabilidades para todos los niveles de la DTI.
- Establecer un método externo imparcial para evaluar el desempeño de los Servicios de TI.

5. Conclusión

Estos resultados nos han servido para monitorear el desempeño y seguir contribuyendo a la mejora continua de los servicios que se brindan a la comunidad universitaria. Actualmente se siguen incrementando servicios a la certificación, así como también se

está realizando un plan de transición para nueva versión de la norma ISO/IEC 20000-1:2018.

Referencias

[1]	Jan van Bon, Ajen de Jong, A. Fundamentos de ITIL. 2008. Volumen 3 (p.15).
[2]	Catálogo de Servicios de la DTI https://www.uanl.mx/utilerias/serviciosdgi/ http://dti.uanl.mx/servicios-en-linea/

Caso de la Universidad del Rosario: Implementación del Sistema Integral de Información

Consuelo Ardila Castro, Yerly Yudibia García Falla

Departamento de Tecnología, Informática y Comunicaciones
Universidad del Rosario, Bogotá – Colombia
consuelo.ardila@urosario.edu.co, yerly.garcia@urosario.edu.co

Resumen. La Universidad del Rosario es una institución de educación superior fundada en el año 1653, ha sido de las instituciones educativas más antiguas de Colombia y durante estos 365 años ha funcionado sin interrupciones, su misión ha sido "Impartir una sólida formación ética, humanística y científica que, unida a la investigación y a una idónea y exigente docencia, permita a esta comunidad educativa formar integralmente personas insignes y actuar en beneficio de la sociedad, con un máximo sentido de responsabilidad"¹ [1] ofrece así carreras de pregrado, postgrado, cursos de educación continua, tiene su propia área de investigación y de proyección social como aporte hacia la comunidad, hoy en día se enfoca en atraer y mantener estudiantes de gran rendimiento académico y ofrecer todo aquello que facilite el aprendizaje y formación integral de los nuevos dirigentes de nuestro país, de ésta manera necesita consolidar la información que califique a sus estudiantes para entender sus necesidades y estar un paso adelante disponiendo los recursos que se requieren, adicionalmente necesita conocer las cifras y los indicadores que generan impacto en la comunidad y que le permitan una toma de decisiones eficiente para garantizar el futuro institucional, es así como se ha creado el sistema integral de información para responder activamente a las múltiples preguntas de negocio y actuar acorde a lo indicado por la información que se procesa.

Palabras Clave: Lago de Datos, ETL, Analítica de Datos, Inteligencia de Negocios, Modelo Tabular, Cubos Multidimensionales.

Eje temático: Smartcampus y tendencias tecnológicas emergentes en la Universidad (aplicaciones de la inteligencia artificial, blockchain, analítica, IoT, impresión 3D, experiencia inmersiva).

¹ <https://www.urosario.edu.co/La-universidad/Mision-y-Vision/>

1. Introducción

La Universidad del Rosario es una institución de educación superior fundada en el año 1653, ha sido de las instituciones educativas más antiguas de Colombia y durante estos 365 años ha funcionado sin interrupciones, su misión ha sido "Impartir una sólida formación ética, humanística y científica que, unida a la investigación y a una idónea y exigente docencia, permita a esta comunidad educativa formar integralmente personas insignes y actuar en beneficio de la sociedad, con un máximo sentido de responsabilidad"², como parte de las funciones sustantivas de la universidad se encuentran las siguientes:

Docencia como función central, la cual está constituida por una serie de procesos sistemáticos enfocados a la formación integral de los estudiantes a través de metodologías pedagógicas activas centradas en el aprendizaje. Estos procesos contienen la planeación, ejecución, evaluación y mejoramiento de las prácticas de enseñanza y aprendizaje, la verificación de los resultados de las actividades curriculares y extracurriculares realizadas para el desarrollo de los estudiantes. Para apoyar estos contextos se han definido estrategias enfocadas en consolidar una planta profesoral de excelencia comprometida con la formación de los estudiantes, ofrecer programas de actualización avanzadas acorde a las exigencias del mercado, incentivar y desarrollar la investigación en el ámbito de la educación superior para proyectar resultados hacia la innovación y el mejoramiento de la gestión académica y de los programas de mejoramiento³. En docencia se ofrecen carreras de pregrado, postgrado en la modalidad presencial y virtual contando con una población estudiantil de 14.951 estudiantes. [2] En educación continua (capacitaciones profesionales a ejecutivos y funcionarios de empresas públicas y privadas para profundizar, actualizar y desarrollar competencias específicas a través de diplomados, cursos, seminarios y talleres) según cifras del boletín estadístico de 2018 se tuvieron 15.130 participantes en 593 eventos⁴[3].

Investigación se cuenta con una división de investigación donde se prioriza el crecimiento y la consolidación científica, para avanzar en el conocimiento que alimente las ciencias, las disciplinas y las profesiones aportando al ejercicio del análisis, la crítica y la construcción social; y contribuir a la proyección de acciones en beneficio del desarrollo humano. Las estrategias definidas para apalancar esta visión han sido promover un entorno favorable para el desarrollo científico consolidando grupos de investigación, diversificar los mecanismos de financiación investigación para hacerla viable y sostenible, fomentar la divulgación de los resultados e investigación, consolidar los semilleros de investigación y los programas de jóvenes investigadores,⁵ es así como en el cierre del 2018 se contaban con 360 proyectos de investigación en ejecución.

Extensión consiste en una serie de actividades donde la universidad integra los saberes, la investigación, la docencia y las necesidades del entorno para proyectarse a la comunidad y contribuir con su desarrollo económico, político, social y cultural, las

² <https://www.urosario.edu.co/La-universidad/Mision-y-Vision/>

³ <https://www.urosario.edu.co/Documentos/La-universidad/PEI.pdf>

⁴ <https://www.urosario.edu.co/Boletin-estadistico/>

⁵ <https://www.urosario.edu.co/Documentos/La-universidad/PEI.pdf>

estrategias definidas para su desarrollo son facilitar la aplicación y difusión del conocimiento generado y desarrollado por la institución, desarrollar la extensión desde las fortalezas, capacidades e intereses y ámbitos de conocimiento de las diferentes unidades académicas, preservar el patrimonio artístico, arquitectónico y cultural para hacer valer la memoria de la institución como parte de la historia del país.⁶ Acorde con las cifras del 2018 se realizaron 100 actividades de extensión y se cuenta con 57 convenios vigentes de prácticas y pasantías.

Como parte de sus objetivos estratégicos la universidad le apuesta a su estrategia de internacionalización, por medio de la promoción de asignaturas en otros idiomas, la apertura de summer schools, la creación y fortalecimiento de convenios con instituciones internacionales para intercambios, pasantías y doble titulación contando así con 344 estudiantes internacionales y 124 profesores internacionales, 184 estudiantes en movilidad saliente y 157 en movilidad entrante.

La universidad cuenta con la acreditación institucional de alta calidad la cual se hace referencia a cómo una institución y sus programas orientan su deber ser hacia un ideal de excelencia, y pueden mostrar alta calidad mediante resultados específicos, tradición consolidada, impacto y reconocimiento social. Este proceso es de carácter voluntario, y consta principalmente de una autoevaluación institucional y una evaluación por pares académicos externos que se centran en una mirada integral a los servicios que las instituciones educativas ofrecen a la sociedad. [4]

Actualmente la universidad se encuentra inmersa ante retos que permitan transformar la educación superior, está enfocada en desarrollar 6 objetivos estratégicos que la consolidaran como la primera opción para aprender en el país, debe preservar lo que ésta institución representa para los colombianos como memoria histórica y consolidación del país, por dichas razones se hace necesario tener métodos para la toma de decisiones eficientes garantizando el futuro de la institución, para lo cual se ha dado relevancia a la información y a los datos institucionales, que acorde a las cifras generadas revelan el estado operacional y estratégico en determinado momento.

De esta forma la universidad ha creado el Sistema Integral de Información para consolidar un modelo de gestión de información empresarial que unifique la información (métricas, indicadores o términos de negocio) integrando desde un punto de vista lógico las fuentes de datos relevantes de la organización, incorporando un conjunto de procesos y herramientas que posibilitan el tratamiento de la información con una visión única que permite la estandarización de conceptos de negocio y garantiza la consistencia de resultados, es decir, una única visión de la realidad empresarial para el soporte de toma de decisiones, realizar mejores predicciones, implementar mejores estrategias y tener mejores tácticas que permitan adelantarse a los acontecimientos, entre otras.

Bajo esta perspectiva la universidad ha crecido en el uso de información corporativa enfrentándose al reto de consolidar una plataforma tecnológica adecuada, que soporte las necesidades de información planteadas por el negocio más aún con un carácter evolutivo en el tiempo.

⁶ <https://www.urosario.edu.co/Documentos/La-universidad/PEI.pdf>

2. Definición del Problema y Planteamiento de la Solución

A mediados los años 2016 se define una visión en el plan integral de desarrollo 2020 y se establecen 6 objetivos estratégicos que ayuden al cumplimiento de la visión.⁷ [5] "En 2020 seremos la primera opción para los estudiantes que buscan aprender a aprender en una comunidad universitaria de alto desempeño y vocación internacional. Movemos fronteras a través de la investigación y la promoción de la equidad y la diversidad. Entregamos a la sociedad los mejores ciudadanos con visión global y somos actores protagónicos de los grandes temas del país." Adicionalmente se tienen los siguientes retos con el propósito de transformar la educación superior:

- Aumento de la productividad.
- Capacidad de respuesta a un entorno cambiante, rendición de cuentas y reducción de costos.
- Mayor control externo por parte del gobierno, sociedad civil y el sector productivo.
- Gran alcance de las nuevas tecnologías que requieren importantes inversiones tanto financieras como de capital humano.
- Masificación de emprendimientos más allá de la educación tradicional que llevan a una innovación constante.
- Aumento de la competencia en el mercado.
- Atracción de profesores, estudiantes y personal de alta calidad.

Para lograr la visión 2020, la atención de los retos anteriormente descritos y contar con información oportuna, de calidad y pertinente para tomar las decisiones correctas y orientar a la Universidad de manera que continúe en su camino hacia el liderazgo y excelencia como un elemento vital en el plan de desarrollo trazado, se crean áreas estratégicas como la División de Planeación y Efectividad Institucional (DPEI), quienes contratan un diagnóstico de una empresa consultora encontrando las siguientes oportunidades de mejora:

- Deficiencias en la integridad de la información.
- Ausencia de políticas y mecanismos para la gobernabilidad de la información.
- Incompleta alineación de la gestión de la información con la estrategia de la Universidad.
- Necesidad de desarrollar una cultura de información que impulse su uso y apropiación.

Desde la Dirección de Planeación y Efectividad Institucional:

- Falta de coordinación en el reporte de información institucional.
- Desconocimiento de la información que las áreas reportan.
- Excesiva descentralización en la recolección y consolidación de la información.

Así mismo para cumplir con el objetivo estratégico "lograr la sostenibilidad de largo plazo sin sacrificar el posicionamiento en el corto plazo" y respondiendo a las opciones de mejora detectadas se decide implementar un sistema integral de información, evaluación y seguimiento que soporte las decisiones estratégicas.

⁷ <https://www.urosario.edu.co/La-universidad/Mision-y-Vision/>

Para crear el sistema integral de información como primera medida se debe crear el comité estratégico de información en el cual se definan las variables de negocio, los indicadores y las mediciones requeridas para dar respuesta a las preguntas de negocio y que al interior de la universidad todos hablen el mismo lenguaje.

3. Comité Estratégico de Información

El comité estratégico de información es la instancia mediante la cual se coordina y articula la gestión estratégica de información en la Universidad del Rosario, está integrado por las divisiones representativas de la institución como se muestra en la siguiente figura.



Fig. 1. (Estructura del comité estratégico de información.)

Dentro de las principales funciones del comité estratégico de información se encuentran las siguientes:

- Establecer los lineamientos de manejo de información y la gestión de calidad de los datos.
- Fomentar la cultura del análisis de información y su uso para la toma de decisiones.
- Apoyar la definición de acuerdos entre áreas para la medición y reporte de información.
- Definir los productos y servicios para la publicación de la información estratégica teniendo en cuenta criterios de privacidad, pertinencia, detalle, público objetivo y canales a utilizar.
- Coordinar las actividades de definición, seguimiento, evaluación y mejoramiento a la implementación de la cadena de valor y proceso de gestión de información.
- Establecer los lineamientos para el desarrollo de auditorías de control y aseguramiento de calidad de la información y liderar el seguimiento a los planes de mejoramiento que se deriven de las mismas.

En el comité estratégico de información también se crea el esquema de gobierno de la información, donde el Departamento de Tecnología, Informática y Comunicaciones actúa como el habilitador que implementa y se apalanca en las tecnologías para soportar la estrategia en gestión de información:



Fig. 2. (Modelo de gestión estratégica de información.)

Una vez establecido el comité estratégico de información se realizan las sesiones de comité, con el objetivo de establecer la definición de las variables de negocio, las métricas, los conteos, los operandos, los niveles de desagregación, que hacen parte de la definición del modelo de indicadores institucionales, generando así la metodología de conteo de indicadores en la universidad. Con este documento el departamento de tecnología realiza el estudio de mercado y selección del proveedor que apoye la creación del sistema integral de información, la consolidación de la arquitectura de información y el establecimiento de modelo empresarial para análisis de información.

4. Sistema Integral de Información

Como método de solución a este gran reto la Universidad del Rosario implementó una solución tecnológica construyendo un repositorio de datos basado en un lago de datos, disolviendo la complejidad en un modelo de capas de información, cada una con un objetivo en específico para consolidar una capa analítica; desde la cual se crea un ambiente colaborativo y para permitir el crecimiento del modelo, así como la reducción de tareas repetitivas, concentrado su mayor esfuerzo en ganar más conocimiento de la universidad y su entorno, para mejorar la calidad de vida de cada persona del ecosistema universitario.

Los principales objetivos de este sistema son los siguientes:

- Configurar una arquitectura de información escalable para crecer el modelo.
- Consolidar las fuentes de datos que tiene la Universidad en un repositorio.
- Desarrollar la lógica de los indicadores definidos por el Comité Estratégico de Información.
- Crear la visualización y formulación de los indicadores estratégicos,

flexible para que la institución pueda modificar y/o crear nuevos indicadores, a partir de los existentes.

- Implementar consultas que permitan generar datos e indicadores para tomar decisiones.
- Crear un modelo de autoconsumo de información y dejar la base constituida para que la universidad pueda crear nuevos tableros.

El desarrollo del proyecto en el tiempo tuvo las siguientes etapas, las cuales generaron los insumos necesarios para su desarrollo:



Fig. 3. (Etapas en las que se desarrolló el sistema integral de información.)

4.1 Modelo Analítico de Capas de Información

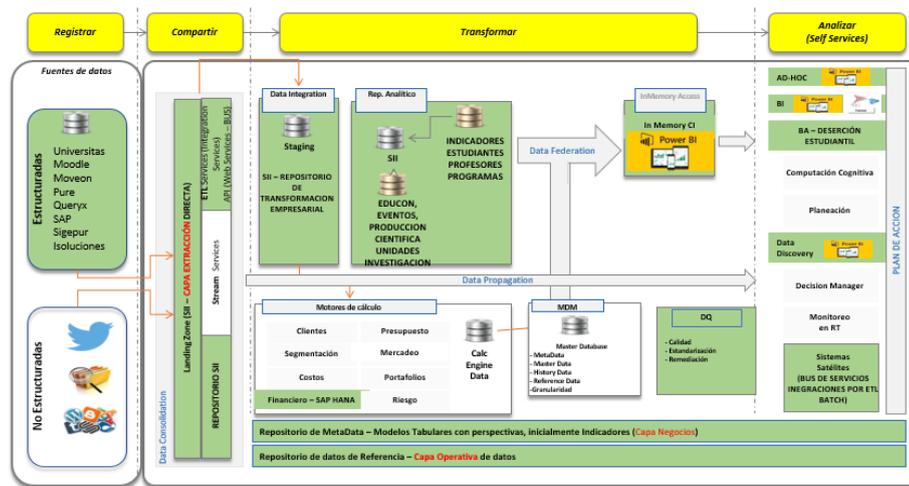


Fig. 4. (Modelo analítico de capas de información.) En este mapa se pueden apreciar las capas de información y los procesos que intervienen en cada una, con el proyecto se desarrollaron los componentes en color verde y se dejaron las bases para continuar creciendo el modelo hacia una analítica avanzada. [6].

Capa 1 - Fuentes Transaccionales: se obtienen los datos brutos provenientes de procesos operativos, los cuales por su naturaleza no brindan la capacidad de transformar dichos datos en información.

Capa 2 - Capa de Extracción: Contiene los datos procedentes de los sistemas

transaccionales, se hace una carga inicial y cargas incrementales a través de procesos de sincronización con periodicidades definidas por el negocio, las particularidades de ésta capa son:

- Quitar la carga del tratamiento de los datos a las fuentes transaccionales. vistas en la capa 1, debido a que no están diseñadas para ese fin.
- Hacer procesos diarios de extracción los cuales tienen como objetivo la actualización constante del modelo.
- Mantener únicamente información del último proceso (diario o mensual)
- Mantener la integridad referencial espejo de la capa 1.
- Hacer sabanas de información para entregar reportes detallados para evitar sobrecarga al sistema transaccional, liberando recursos los cuales hacen más óptimos los procesos operativos, dichos reportes mantienen únicamente la información del último proceso.
- Reportar errores automáticos en caso de inconsistencias.

Capa 3 - Capa de Transformación: Se efectúan procesos de limpieza de datos, manteniendo la integridad referencial y también se ajustan atributos enfocándolos al modelo de negocio, los cuales van a ser más amigables con funcionarios analíticos. Las particularidades de ésta capa son:

- Se comunica únicamente con la capa 1 (fuentes transaccionales) y realiza procesos que son altos en consumo de recursos como memoria, procesador y escritura en disco.
- Mantiene únicamente información del último proceso (diario o mensual).
- Realiza procesos de homologación y consolidación de la información.
- Reporta errores automáticos en caso de inconsistencias.
- Genera reportes de auditoría del tratamiento de los datos.

Capa 4 - Capa de Negocio: Consolida la información recolectada diariamente alimentándola durante diferentes periodos, aplicándole tratamientos con las buenas prácticas definidas en gestión de información, realizando agrupamientos y aplicación de reglas de negocio, los cuales permiten mantener su historia y dejarla en su punto para posteriormente aplicar analítica. Sus particularidades son:

- Separación de los datos en barriles de información puesto que en algunos casos no se requiere analizar toda la historia si no algún periodo de tiempo determinado.
- Se pueden generar sabanas de información, pero tiene la ventaja de poseer datos históricos y al estar optimizada para la lectura de grandes volúmenes de datos lo genera sin inconvenientes
- Está enfocada en temáticas de la universidad; por ejemplo, estudiantes, profesores, información financiera entre otros.
- Se encuentra estandarizada en un lenguaje único para todas las áreas de la universidad. Acorde a las definiciones del comité estratégico de información, con esto cuando se menciona un atributo significa lo mismo en la institución.

Capa 5 - Capa de Explotación: Permite que los procesos analíticos de la universidad puedan usar los datos de la capa 4 (de negocio), transformarlos en información y de esta información generar conocimiento, el cual es usado para generar estrategias, tácticas, acciones inmediatas y de esta forma mantener un bienestar en el

ecosistema universitario trabajando así en sinergia con las áreas de la universidad y estudiantes. Sus particularidades son:

- Se usan modelos tabulares y multidimensionales los cuales están enfocados en analizar los diferentes hechos que ocurren cada día desde diferentes puntos de vista; obteniendo patrones ocultos entre la información.
- Se utiliza Power BI como herramienta de análisis interactiva, así como también para la presentación de resultados, debido a que con simples movimientos presenta lo que está ocurriendo en el momento en que el usuario así lo demande, por ejemplo y como caso de éxito se presenta el seguimiento de los diferentes indicadores de cada área, los cuales muestran el comportamiento de la universidad desde diferentes ángulos.

5. Terminología

Lago de Datos: Consiste en un depósito de datos en bruto procedente de los sistemas fuente, conservando su formato y su estructura sin modificaciones, para nuestro proyecto se asocia a la capa de aterrizaje. Generalmente está compuesto por datos estructurados, datos semiestructurados y datos no estructurados.

ETL: Proceso que permite extraer datos de sus fuentes originales, transformarlos acorde a las necesidades del negocio y cargarlos a otras bases de datos.

Analítica de Datos: La analítica de datos es el uso de diferentes herramientas tecnológicas para el descubrimiento, la interpretación, la socialización de patrones significativos en los datos y el análisis de esos patrones para la toma de decisiones para lograr el estado deseado de la organización.

Inteligencia de Negocios: Procesos mediante los cuales una organización transforma datos en información, la información en conocimiento y ese a su vez se utiliza para crear estrategias y tácticas las cuales permitan tomar acciones en un momento determinado.

Modelo Tabular: Son bases de datos que se ejecutan en memoria, ofrece un enfoque de modelado relacional y su almacenamiento se hace en base a columnas y no en registros. [7]

Cubos Multidimensionales: Son bases de datos que permiten almacenar grandes volúmenes de información a través de medidas y dimensiones; mejorando la eficiencia, la velocidad de respuesta y la simplicidad de las consultas. Su almacenamiento se realiza en base a registros y no en columnas. [8]

6. Resultados Obtenidos

Creación de un ente regulador de la información al interior de la Universidad: el Comité Estratégico de Información es el que da los lineamientos en gestión de información al interior de la institución, realiza sesiones cada dos meses, analizando las nuevas necesidades y planteando estrategias conjuntas de solución.

Generación de información institucional automática: Se ha entregado un modelo de indicadores estratégicos, los cuales son calculados automáticamente por el sistema

acorde a las fechas de corte y se disponen en una plataforma donde se tiene acceso inmediato a las cifras, así mismo se pueden generar los soportes de las cifras generadas en sus niveles de desagregación.

Este proceso ha disminuido a la carga operativa en un 80%, en cuanto antes los cálculos de los indicadores estratégicos se realizaban de manera manual. Esto ha mejorado la calidad de vida de los colaboradores debido a que el enfoque tiende a ser menos operativo y mayor análisis de datos.

El sistema también brinda la posibilidad de ver la evolución de las cifras a través de diferentes periodos, permitiendo tomar decisiones y observar el comportamiento de la gestión a través de los años.

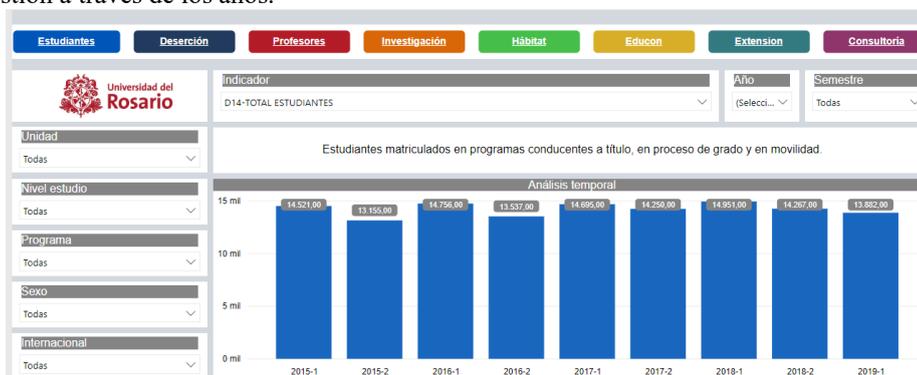


Fig. 5. (Tablero de seguimiento de cifras generado por el sistema integral de información).

Agilidad en la generación de reportes por demanda: Se ha entregado el modelo de autoconsumo de información, el cual facilita a los usuarios la generación de sus propios reportes, acorde a las diferentes perspectivas configuradas, las cuales heredan las reglas de negocio y permiten generar salidas de información consistentes con las cifras estratégicas.

Crear nuevas plantillas de indicadores, cuadros de mando, reportes con la información base de los indicadores estratégicos.

Transparencia y confianza en las cifras generadas: al haber un único ente que regula la información al interior de la universidad, al haber establecido una metodología de cálculo conjunta y al disponer las cifras y su origen para todos los usuarios ha generado la confianza y la certeza de que las cifras que se reportan son oficialmente válidas.

Construcción de una cultura institucional de análisis de información: La orientación hacia los empleados de la universidad es a enfocarse en el análisis de los datos y no en la operatividad que tenían antes para reportar las cifras.

7. Conclusiones

La Gestión de información permite el establecimiento de un modelo empresarial de información, el cual que unifica la información de la universidad (métricas, indicadores o términos de negocio) integrando desde un punto de vista lógico las fuentes de datos relevantes de la organización, incorporando un conjunto de procesos y herramientas

que posibilitan el tratamiento de la información con una visión única que permite la estandarización de conceptos de negocio y garantiza la consistencia de resultados, es decir, una única visión de la realidad empresarial para el soporte de toma de decisiones, realizar mejores predicciones, implementar mejores estrategias.

Las instituciones que han crecido en el uso de información corporativa se enfrentan al reto de tener plataformas tecnológicas adecuadas que soporten las necesidades de información planteadas por el negocio, para enfrentar estos retos es necesario trazar una ruta de crecimiento adecuada que cubra unos frentes principales de forma escalable y defina políticas y lineamientos perdurables en el tiempo en un marco de gobierno de datos.

8. Próximos Pasos

En camino está un proyecto de gobierno y calidad de datos basado en las buenas prácticas establecidas por el DAMA – Data Management. [9]

Crece el sistema integral de información integrando nuevas fuentes de información y generando nuevos indicadores.

Establecer modelos predictivos en la universidad a partir de fuentes estructuradas y no estructuradas.

9. Referencias

1. Universidad del Rosario, [en Línea]. Available <https://www.urosario.edu.co/La-universidad/Mision-y-Vision/>. [Último acceso: 15 de mayo de 2019].
2. Universidad del Rosario, [en Línea]. Available <https://www.urosario.edu.co/La-universidad/Mision-y-Vision/>. [Último acceso: 15 de mayo de 2019].
3. Universidad del Rosario, [en Línea]. Available <https://www.urosario.edu.co/Boletin-estadistico/>. [Último acceso: 31 de marzo de 2019].
4. Universidad del Rosario, [en Línea]. Available <https://www.urosario.edu.co/Acreditacion-Institucional/inicio/>. [Último acceso: 21 de abril de 2019].
5. Universidad del Rosario, [en Línea]. Available <https://www.urosario.edu.co/La-universidad/Mision-y-Vision/>. [Último acceso: 15 de mayo de 2019].
6. Presentación propuesta comercial "Sistema de Información Integral V12.pdf" u RoadMap UR. PPT página 11.
7. Microsoft [en Línea]. Available <https://docs.microsoft.com/es-es/sql/analysis-services/tabular-models/tabular-models-ssas?view=sql-server-2017>. [Último acceso: 10 de mayo de 2019].
8. Microsoft [en Línea]. Available <https://docs.microsoft.com/es-es/system-center/scsm/olap-cubes-overview?view=sc-sm-2019>. [Último acceso: 10 de mayo de 2019].
9. Dama-Dmbok: Data Management Body of Knowledge (libro en inglés), Edición 2 2017

Estrategia de implementación para los sistemas de gestión de servicios de TI: Data Center Infrastructure Management y SCG

Yazmín Diana Reyes Torres
yazdia@unam.mx

Resumen. La Universidad Nacional Autónoma de México (UNAM) cuenta con 356,530 alumnos, 41,318 académicos y presencia internacional. Los servicios de Tecnología de Información (TI) que se requieren abarcan un amplio espectro aunado a la velocidad en la que avanza la tecnológica que hace se vuelva un sistema complejo. La Dirección General de Cómputo y de Tecnologías de Información y Comunicación (DGTIC) de la UNAM a través de su Dirección de Telecomunicaciones (DT) ha tomado acciones que permiten aproximarse a la eficiencia y eficacia en el otorgamiento de estos servicios de TI.

La gestión de la tecnología en uno de sus enfoques que está directamente relacionada con la cualificación de las personas encargadas de desarrollar o ejecutar los procesos de misión crítica en las organizaciones que conlleva la atención en la gestión del conocimiento como una de las principales misiones de las Universidades hoy en día, y para la DGTIC es uno de sus retos diarios.

Es identificable la experiencia tecnológica de las áreas de TI, que con una debida gestión puede permitirles contribuir en la toma de decisiones e impactar en la operación diaria de incidentes y de proyectos innovadores. La Dirección de Telecomunicaciones a través del Centro de Atención a Usuarios (CAU) de la DGTIC se ha destacado en este sentido ya que sus propuestas han redundado en mejoras prácticas para la entrega de los servicios que se brindan a la comunidad universitaria. Buscando la mejora continua y en el afán de colaborar con el logro de los objetivos de la UNAM, la participación del CAU ha resultado estratégica y ha sido un parteaguas en la manera de gestionar servicios, recursos y tecnología.

La finalidad de este documento es presentar la estrategia de implementación de gestión de servicios de TI mediante las herramientas Data Center Infrastructure Management (DCIM), Sistema de Control de Gestión (SCG) y estamos en el proceso de la implementación del sistema de Gestión Libre del Parque Informático (GLPI).

Palabras Clave: DCIM, SCG, GLPI, gestión del conocimiento, experiencia tecnológica, mejora continua.

Eje temático: Transformando los servicios hacia la mejora de la experiencia del usuario (comunidad universitaria).

1. Introducción

A finales del 2016 se presenta una falla en el SITE principal del área de TIC que provocó que se dejaran de brindar servicios en las bases de datos, sitios web, correo electrónico, etc. Además de la pérdida de información no se contaba con una visión integral del centro de datos. La recuperación fue ineficiente.

El requerimiento surge a partir de este incidente donde se detectan áreas de mejora, principalmente en la gestión de la tecnología y del conocimiento, que a su vez contribuyeran en la toma de decisiones, acorde a las mejores prácticas.

Data Center Infrastructure Management (DCIM)

La falta de una buena gestión tecnológica del centro de datos era necesaria, de una gestión del conocimiento⁸, así como de un inventario disponible de la infraestructura instalada a nivel de servidores, máquinas virtuales, cableado eléctrico, switches, capacidad de consumo eléctrico, enfriamiento por usar y servicios.

Se partió de dos posibles soluciones: diseño de un nuevo sistema o buscar algún sistema de software desarrollado que pudiera agilizar todo el proceso de desarrollo de una nueva aplicación. A partir de ahí empieza la búsqueda en la red sobre soluciones ya existentes, que ayudaran a subsanar la problemática que se había planteado, para ello se identificaron tendencias, casos de éxito que otros ya habían probado y que se pudieran ser acordes a las mejores prácticas de TI, de este ejercicio apareció un concepto denominado Data Center Infrastructure Management (DCIM) como una solución. Un software que permite proporcionar una vista integral del centro de datos a nivel de servidor, racks, red, y potencia.

Se selecciona uno de ellos y se realiza la instalación, se prepara un prototipo, y se presenta ante los responsables del Centro de datos (el jefe de servidores, el jefe del servicio electromecánico, el jefe de red y los directores involucrados, así como con el director general). En la presentación los directores aprueban la solución y así nace este proyecto en diciembre del 2016.

Sistema de Control de Gestión (SCG)

La gestión de las solicitudes por escrito (oficios) que ingresaba a la dirección general estaba a cargo del personal de la misma dirección y no contaba con un control unificado que permitiera conocer el estado de la solicitud, el tiempo de respuesta, así como dar seguimiento y asignarlo, para la atención, a varias áreas de la organización, finalmente no se contaba con un repositorio digital donde se resguardara la gestión de estos documentos.

Este sistema se crea para gestionar adecuadamente las solicitudes de las dependencias universitarias hechas a la dirección de TI, en el entendido que es

⁸ El conocimiento es una forma de alto valor de la información que está lista para ser aplicada a decisiones y acciones (Davenport, De Long y Beers, 1998)

información de carácter estratégico para la universidad y amerita una gestión pronta y expedita. Por lo que el principal requerimiento es implementar mecanismos de seguimiento y un tablero de control donde se pueda visualizar de una forma muy sencilla, la cantidad de pendientes que existen por área y consultar los pasos que se han estado llevando a cabo para atender las diversas solicitudes de la comunidad universitaria y que además permita el trabajo colaborativo entre las diferentes direcciones de la DGTIC con la finalidad de que la atención se realice eficientemente y sin periodos de tiempo muertos. La puesta en marcha se inicia en septiembre del 2017.

2. Descripción de la solución

Para ambos casos nos ayudamos de dos modelos, el primero es el de las funciones principales de la gestión tecnológica [1] y la segunda basada en el proceso de mejora continua del servicio de ITIL. Adaptándolos a nuestras propias necesidades y contexto.

Tabla 1. Funciones principales de la gestión tecnológica.

Función	Modelo Solleiro	Adecuación del modelo para la DGTIC
Inventariar	<p>Recopilar tecnología disponible en el área</p> <p>Conocer las tecnologías utilizadas y dominadas que constituyen su patrimonio tecnológico</p>	<p>En este sentido cuando se plantean nuevos retos, ya sea por nuevos servicios o cuando se presenta un incidente, lo primero a realizar es conocer con lo que contamos, que información tenemos y quien la administra.</p>
Vigilar	<p>Alertar sobre la evolución de nuevas tecnologías</p> <p>Sistematizar las fuentes de información</p> <p>Vigilar la tecnología de las otras áreas</p> <p>Identificar el impacto posible de la evolución tecnológica sobre la actividad del área</p>	<p>Estar alerta sobre lo que están utilizando otras organizaciones y conociendo sus experiencias nos ha servido para contar con los elementos necesarios para determinar los pasos a seguir, saber los casos de éxito de otros nos permite actuar con rapidez al buscar solucionar problemas o plantear nuevas propuestas.</p>
Evaluar	<p>Determinar el potencial tecnológico propio</p> <p>Estudiar posibles estrategias para la evaluación de soluciones</p>	<p>Ver quien puede tener las capacidades y aptitudes necesarias dentro de la dependencia.</p> <p>Identificar un pool de software o soluciones posibles</p>

	Identificar posibles alianzas tecnológicas	Identificar de estas cual es viable de implantar en nuestra organización, porque hay que saber para cual estamos preparados y para cuáles no.
Enriquecer	<p>Diseñar estrategias de investigación y desarrollo</p> <p>Priorizar tecnologías emergentes, clave y periféricas</p> <p>Definir una estrategia de adquisición de equipo y tecnologías externas</p> <p>Establecer proyectos conjuntos y alianzas</p> <p>Determinar estrategias de financiamiento a proyectos</p>	<p>Dividir el problema en pequeñas piezas para empezar a revisarlas una a una y determinar si podemos desarrollar algo internamente o si es preferible optar por alguna solución que ya hay en el mercado.</p> <p>Que el software elegido pueda ser modificado o integrado a otros productos ya existentes dentro de la organización.</p> <p>Poder agregar nuevas funciones que no se incluyan en la solución original.</p> <p>Seleccionar el equipo de trabajo que por su perfil nos pueda garantizar una solución viable en el menor tiempo posible, conseguir patrocinadores de la alta dirección para garantizar el éxito del proyecto.</p>
Asimilar	<p>Explotación sistemática del potencial tecnológico mediante:</p> <ul style="list-style-type: none"> • Documentación de tecnologías de las universidades y otras organizaciones • Desarrollo de aplicaciones derivadas de tecnologías genéricas • Gestión eficiente de recursos 	<p>Adaptar o modificar características solicitadas a la tecnología en la medida de lo posible.</p> <p>Empezar a usar la tecnología pieza a pieza, característica a característica y ver que partes nos dan un mayor beneficio.</p> <p>Gestionar el uso de la tecnología seleccionada.</p>
Proteger	Establecimiento de la política de propiedad intelectual	Gestionar el conocimiento obtenido, para la toma de decisiones futuras a través de una eficiente documentación de los proyectos.

Para ITIL la fase de Mejora Continua (fig.1), se trata de identificar e implementar las mejoras necesarias para garantizar que los servicios de tecnología sean más efectivos y eficientes además de que estén alineados con las cambiantes necesidades del negocio.

ITIL define la gestión de servicios de tecnología como una capacidad organizacional que le permite a una compañía administrar sus servicios de tecnología, para que estos sean capaces de soportar los procesos de negocio de la organización y generarle valor real al negocio.



Fig.1. Mejora continua. <http://inventarios.org/2016/03/04/los-7-pasos-de-la-mejora-de-procesos-definidos-en-itol/>

En este sentido nunca perdemos de vista estos propósitos; que la tecnología que proporcionemos a la universidad ayude a generar valor y que sea pueda mejorar continuamente.

¿Por qué DCIM?

La gestión tecnológica es la vía óptima para combinar recursos humanos, técnicos y financieros para el cumplimiento de los objetivos de la organización (Solleiro y Castañón, 2008).

La administración de la infraestructura de un Data Center es una tarea compleja, y en el mercado existen herramientas llamadas **DCIM, Data Center Infrastructure Management** que proveen la capacidad de lograr un buen desempeño de los recursos administrados, buscando la optimización y planificación continua de una manera simple, considerando que cada vez hay más interdependencia entre la capa física y la capa lógica de la infraestructura en un centro de datos. Existen muchas soluciones DCIM comerciales, pero uno de los requerimientos principales era una solución rápida y de bajo costo que permitiera una optimización, otro punto importante es la falta de presupuesto por lo que la búsqueda se centró en el ámbito del software de código abierto ya que algunas soluciones comerciales dividen su costo en un software central y un costo extra por nodo a administrar. OpenDCIM está bajo la licencia GPL v3.

Se elige openDCIM por ser gratuita. Según el sitio web <https://www.opendcim.org/> el objetivo número uno de openDCIM es eliminar la excusa para que cualquiera pueda realizar un seguimiento del inventario de su centro de datos.

OpenDCIM permite tener el inventario físico de los activos (incluyendo su ubicación) de los centros de datos (Fig.2), así como administrar los espacios, la energía eléctrica y condiciones ambientales (aire acondicionado). También permite que los administradores tengan a la vista la configuración de red de los dispositivos, saber la capacidad actual de cada rack, cuántas tarjetas de red tiene cada dispositivo y a que switch está conectado, cuanta potencia se requiere y cuanto está consumiendo por rack, permite conocer las conexiones de cableado entre los gabinetes y entre dispositivos.

UNAM Universidad Nacional Autónoma de México

Centro de Datos Servidores
Estadísticas del CPD

Buscar por Nombre

[Avanzado]

Informes

- Formulario de Solicitud de Rack
- Administración de Usuarios
- Escalados
- Project Catalog
- Gestión de Plantillas
- Gestión de las Infraestructuras
- Bulk Importer
- Gestión de Suministros
- Gestión de la Energía
- Caminos de Conexiones
- Editar Configuración

Servidores (Estado: OK)			
Infraestructura	Ocupado	Sitio	Disponible
Us Totales 589	283	68	566
Porcentaje	28.6%	8.9%	57.1%
Potencia Calculada (W)	6342 Vatios		
Potencia Medida (W)	30010 Vatios		
Potencia Máxima de Diseño (W)	0 W		
Cálculo de la BTU a partir de Vatios Calculados	2042 BTU		
Tamaño del CPD	100 Metros Cuadrados		
Vatios por Metro Cuadrado	83 Vatios		
Total Cabinets	24		
Tarjetas de Refrigeración Mínimas Requeridas (a partir de Vatios Calculados)	2 Tarjetas		
Temperatura Media	0 °F		
Humedad Media	0 %		
PCI Porcentaje bajo (sobrecalentamiento)	0 %		
PCI Porcentaje alto (Riesgo satelador)	100 %		

Fig. 2 Sistema OpenDCIM.

Sistema de Control de Gestión (SCG)

La Dirección de Telecomunicaciones de la DGTIC hace más de 15 años que cuenta con un sistema de tracking que ha permitido llevar el seguimiento de las solicitudes de servicios e incidentes en materia de telecomunicaciones que demanda la Comunidad Universitaria, como correo electrónico, redes, telefonía, videoconferencias, etc. Permitiendo la gestión desde su puesta en marcha de más de 143.000 tickets, este sistema en su momento tuvo un costo de \$6000.00 pesos de inversión que a la universidad hasta el día de hoy le ha reeditado mucho más de lo esperado ya que se cuenta con acceso al código lo que ha permitido adecuarlo a las necesidades que se han presentado. Por lo que hacer una versión nueva con los requerimientos solicitados por la alta dirección no resulto muy difícil (Fig.3).

The screenshot shows a web-based form for capturing system data. The form is organized into several sections:

- Título (Reporte):** A text input field containing "red abajo".
- Dependencia:** A dropdown menu with "UNAM DEPENDENCIAS" selected. A list of options is visible, including "DIRECCION GENERAL PARA LA PREVENCIÓN Y MEJORA DE LA GESTIÓN INSTITUCIONAL", "ESCUELA NACIONAL DE ENFERMERÍA Y OBSTETRICIA", "ESCUELA NACIONAL DE ESTUDIOS SUPERIORES UNIDAD JURIQUILLA, QUERÉTARO", "ESCUELA NACIONAL DE ESTUDIOS SUPERIORES UNIDAD MÉRIDA", "ESCUELA NACIONAL DE ESTUDIOS SUPERIORES, UNIDAD LEÓN, GUANAJUATO", "ESCUELA NACIONAL DE ESTUDIOS SUPERIORES, UNIDAD MORELIA, MICHOACÁN", "ESCUELA NACIONAL DE LENGUAS, LINGÜÍSTICA Y TRADUCCIÓN", and "ESCUELA NACIONAL DE TRABAJO SOCIAL".
- Áreas DGTIC:** A dropdown menu with "Dirección de Telecomunicaciones" selected.
- Servicio:** A section with three dropdown menus: "Macro Servicio" (Infraestructura de cómputo y telecomunicaciones), "Servicio" (Red UNAM), and "Subservicio" (Conexión a Internet y RedUNAM).
- Datos:** A section with three fields: "Oficio Recibido" (empty), "Asignado A:" (Miriam Alejandra Pérez Corona), and "Prioridad:" (2-Normal).
- Acción a seguir:** A list of checkboxes for actions: "Hacerse cargo de este asunto", "Preparar respuesta para firmar", "Para su conocimiento", "Conciliar con otra área", "Enviar propuesta/respuesta", "Para difusión", "Para firma o rúbrica", "Referir antecedentes", "Tomar nota y archivar", and "Tratar el asunto en próximo acuerdo".
- Adjuntar archivo:** A section with two "Referencia" fields (1 and 2, both Max. 5Mb) and a "Seleccionar archivo" button. A message "No se eligió archivo" is displayed.

Fig. 3 Pantalla de captura del sistema de Control de Gestión.

3. Aspectos críticos y relevantes que resaltar y detallar.

Al intentar implementar una solución tecnológica siempre nos enfrentamos a una resistencia al cambio, la falta de participación y de compromiso por parte de los colaboradores con los cambios propuestos es decisivo en proyectos de este tipo, por lo que en ambas implementaciones fue fundamental contar con el respaldo de la alta dirección.

Otro aspecto crítico fue hacer un levantamiento correcto de los requerimientos, y así favorecer una base firme en la arquitectura, el diseño y el desarrollo de la solución; pruebas basadas en los supuestos del usuario; controlar el crecimiento de los requerimientos. Esto lo logramos al identificar a los actores claves en cada proceso durante la entrega de los servicios, para garantizar la aceptación de las nuevas plataformas. Nos referimos a la idea de hacer del usuario también el proyecto, y que se identificaran los beneficios que ellos obtendrían de él.

El cambiar de paradigma tecnológico se ha vuelto más común, pues después de un periodo de tiempo determinado este es sustituido por uno nuevo, lo que provoca grandes cambios en la organización y ahora mismo estamos en la implementación de un nuevo sistema GLPI que será el futuro de estos dos que ya implementamos; ahora la organización ya ha visto los beneficios obtenidos de estos y está preparada para los próximos cambios por venir.

4. Resultados obtenidos y su impacto

Uno de los principales resultados esperados en el caso del **DCIM** era poder tener acceso a toda la información en conjunto (Fig.4) de las diferentes áreas que intervienen en la administración del Centro de Datos, y en caso de contingencia identificar claramente las afectaciones y las posibles causas de fallas esto se logró satisfactoriamente. Ahora las áreas pueden saber exactamente los requerimientos, o cambios en tiempo y forma y sin lugar a duda saber si se cuanta con los recursos necesarios para atender las necesidades de la comunidad universitaria.

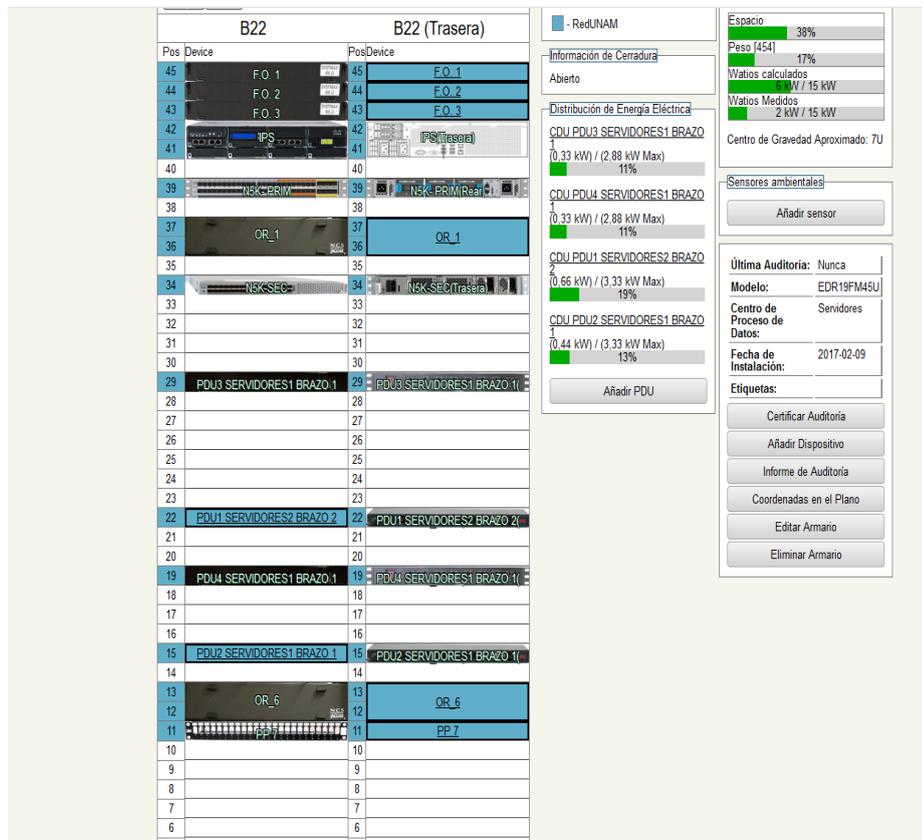


Fig. 4 Equipos alojados en el centro de Datos

En el caso del **Sistema de Control de Gestión** se cumplieron cada uno de los objetivos y nuestro director no pierde de vista la gestión de cada una de las solicitudes que le llegan directamente a él, cuenta con una serie de reportes que le permiten saber el estatus de cada uno de ellos, saber que direcciones han colaborado en la atención de la solicitud, si alguien está retrasando la entrega, cual fue la respuesta que se le dio a estas solicitudes, estadísticas por dependencia,(Fig.5 y 6) etc.

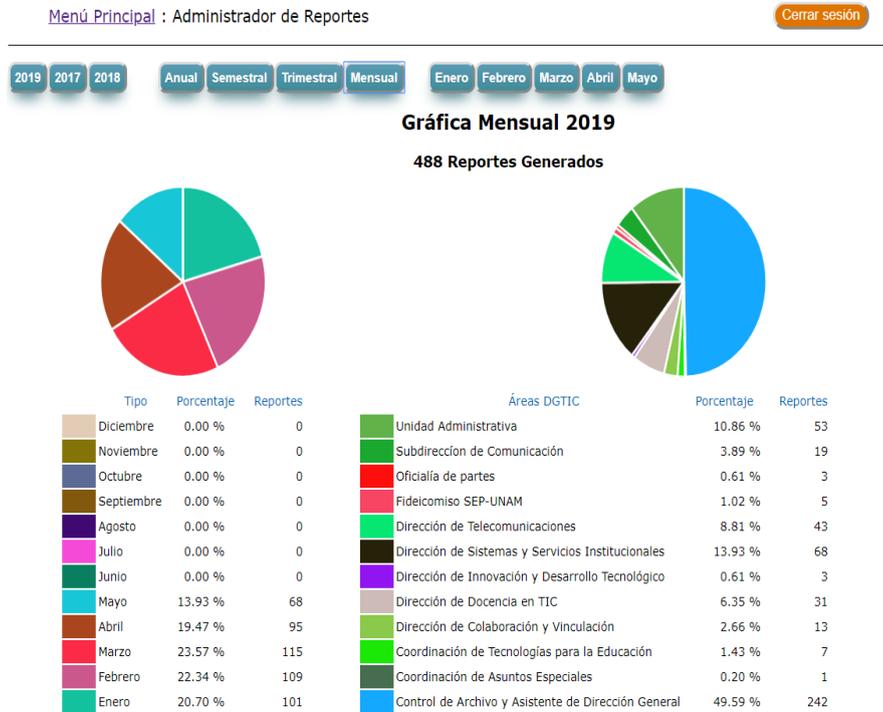


Fig.5 Tablero de control del sistema SCG.

19-0471 Asignar | Cerrar | Editar | Dividir Reporte | Borrar |

Título (Reporte): Cuenta de correo electrónico **Reportes Asociados:**

Dependencia que reporta: Facultad de Estudios Superiores "Zaragoza"

Dirección: Dirección de Telecomunicaciones **Acción a seguir:**
* Hacerse cargo de este asunto

Servicio: Otro

Datos:

Oficio Recibido:	FESZ/DIR/371/19	Estatus:	Abierto
Prioridad:	2-Normal	Enviado a:	
Enviado por:	elvirab	Enviado a las:	21/05/2019 11:55a. m.
Asignado a:	yazdiare	Asignado a las:	21/05/2019 05:46p. m.
Asignado por:	gabrielar	Cerrado por:	
Cerrado por:		Cerrado a las:	

Archivos adjuntos:

- # Archivo: 5195
- Nombre: 19-0471.pdf
- Tamaño: 271966
- Agregado el: 21/05/2019 11:55a. m.
- Command
- Delete

Fig.6 Vista de un reporte del sistema SCG.

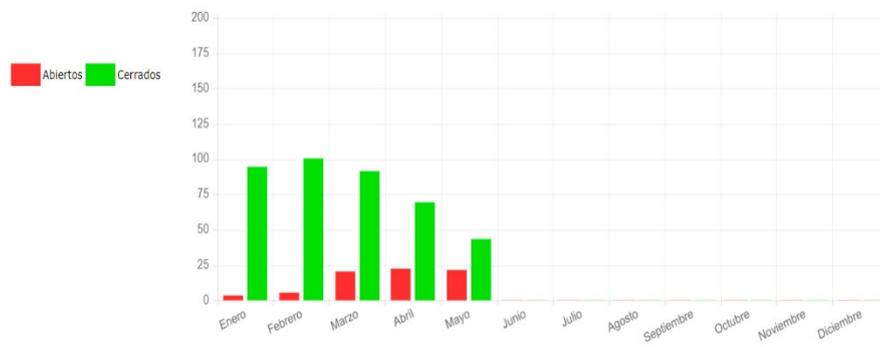


Fig.7 Vista de reportes abiertos y cerrados del sistema SCG.

Aprendizajes

Aprendimos que al cambio de paradigma le antecede la resistencia al cambio pues en primera instancia cuesta trabajo aceptar que hay otras maneras de hacer el trabajo, y ver de manera positiva el esfuerzo que se tiene que realizar para llevar a término estas nuevas iniciativas.

También que la vigilancia tecnológica es y será de suma importancia para poder identificar diversas opciones para las necesidades a las que se enfrentan las Universidades hoy en día.

La gestión de la información ha comenzado a dar sus frutos transformándose en gestión del conocimiento que les da la oportunidad de mejor toma de decisiones a nivel operativo y gerencial.

Un gran acierto fue contar con el respaldo de la alta dirección lo que ayudó a que las tareas fluyeran y se obtuvieran los datos necesarios para llenar las bases de datos del sistema.

Hoy en día es relativamente sencillo el tener el conocimiento de cuál es la carga que se está usando en cada línea de energía eléctrica, evitando accidentes de que al conectar más equipos de los soportados por esta unidad se apaguen todos los equipos conectados a la misma.

El principal acierto es el echar mano de la reingeniería, usar como base el sistema que se tenía y adaptarlo a las nuevas necesidades emergentes, acortando los tiempos de respuesta presentando una solución eficiente y efectiva para la alta dirección.

Tener como patrocinador al director general facilitó enormemente la implementación aun con la resistencia al cambio de algunas áreas.

Aprendimos que es fundamental validar los requerimientos con la alta dirección, y con la gente que opera logrando una armonía que nos garantice la usabilidad de la aplicación.

Impactos no esperados

Gestión de riesgos: a la fecha existe un comité encargado de la buena gestión del centro de datos, con el sistema DCIM logramos darle toda la información que requieren para la oportuna detección de riesgos y con esto poder realizar una buena planeación y revisión con la información fidedigna con la que se cuenta.

Gestión del conocimiento: el comité tiene una impresión completa de todo lo que existe en el centro de datos, pudiendo emitir informes acerca del estado de todos los elementos que lo conforman.

Gestión de la documentación: adicional al sistema DCIM se les generó un módulo dentro de la herramienta de tracking para que se documente cada incidente del centro de datos. Así queda registrado cada uno de los cambios que se estén haciendo en el centro de datos.

Gestión de riesgos: Se han logrado mitigar temas como los cuellos de botella, la falta de comunicación y subsanar la falta de control.

Gestión del conocimiento: Se tiene identificado, temas como cargas de trabajo por área, tiempo de respuesta, solicitudes atendidas por cada entidad universitaria, teniendo los elementos necesarios para la toma de decisiones.

Gestión de la documentación: Se tiene control de cada documento que llega a la dirección general tenido una visión global y expedita que se requiere.

Gestión de los problemas: El sistema está diseñado precisamente para no tener problemas con las solicitudes o detectarlos de forma oportuna para darles solución.

5. Conclusiones

Factores como la globalización económica y los desarrollos en las nuevas tecnologías de información y comunicación, son factores para el aumento en el nivel de competitividad entre las organizaciones, las universidades públicas estamos comprometidas hacer más con menos, en este sentido una buena gestión de la tecnología ayudara en la toma de decisiones.

El aumento de innovación en aspectos estratégicos y operativos ayudara a contribuir en el cumplimiento de los objetivos claves de nuestra universidad. Estar atentos a nuestros competidores, aprender de ellos, vigilar las tendencias tecnológicas, las capacidades de las organizaciones nos permitirá aportar con soluciones adecuadas a nuestra organización y que además le genere valor.

Apoyar la toma de decisiones de manera informada es uno de los principales objetivos alcanzados con la implementación de estas herramientas tecnológicas, mejorando así la entrega de los servicios a la Comunidad Universitaria. Además de apoyar en la administración eficiente de los recursos tecnológicos es una pieza clave para apuntalar los cimientos de gobierno de TI.

Referencias

1. Centro Nacional de Información Biotecnológica, <http://www.ncbi.nlm.nih.gov>
2. Davenport, T., De Long, D Beers, M (1988) "Successful Knowledge Management Projects". Sloan Management REVIEW, 39(2):43-57.
3. Solleiro, J.L y Castañón, R (2008). Gestión Tecnológica Conceptos y prácticas.

Capítulo 4

Tecnologías en la Innovación -Enseñanza e Investigación

Implementación de un modelo de educación virtual predictiva que evita el fracaso asociado a bajos promedios de calificación

Arvey Esteban Granada Aguirre, Cristian Camilo Carmona Gallego^a, Herman Alonso Parra Álzate^b

Grupo de Telesalud, Facultad de Ciencias Para la Salud - Universidad de Caldas, Cra. 25 No 48 – 57
(Manizales, Colombia).

esteban.granada@ucaldas.edu.co, cristian.carmona_g@ucaldas.edu.co,
eduvirtual.telesalud@ucaldas.edu.co

Resumen. Telesalud Universidad de Caldas inició su proceso de formación virtual 10 años atrás bajo un modelo tecno - andragógico, el cual ha tenido diferentes fases de evolución enfocadas en capacitar a profesionales del área de la salud, estudiantes de últimos semestres de carreras a fines al área de la salud y comunidad en general, que por su naturaleza de ocupación, localización geográfica, exigencias de su profesión y estilos de vida, se convierte en un grupo objetivo con necesidades obligatorias de actualización, capacitación y certificación. Es por esta razón, que Telesalud realiza permanentemente una serie de desarrollos tecnológicos basados en inteligencia artificial y analítica de datos que permiten predecir qué perfil de participantes tienen más probabilidades de culminar las capacitaciones cumpliendo con los porcentajes de calificación requeridos para lograr la certificación, y en que momento se deben activar los protocolos de alarma y rescate del estudiante a través de un sistema de acompañamiento que permite predecir de una manera oportuna el comportamiento de los estudiantes y anticipar acciones correctivas ante calificaciones que generen riesgo de no certificación, con el fin de garantizar la culminación de las capacitaciones y obtener los conocimientos requeridos para el buen ejercicio de su profesión.

El grupo de Telesalud y su modelo de educación virtual predictiva identifica patrones de comportamiento y alerta oportuna a partir de un algoritmo que clasifica y mide la evolución de las calificaciones de los participantes al interior de las aulas virtuales, en los diferentes módulos de aprendizaje, activando la secuencia de notificación y recuperación a través de un protocolo permitiendo dar paso al inicio de toma de acciones oportunas de respuesta previo al riesgo de no lograr la certificación.

Palabras Clave: Educación virtual – inteligencia artificial - analítica de datos - modelo de educación – baja deserción – capacitación emotiva – virtualidad - tecnoandragógico - comunidades de aprendizaje – predictiva – predicción - experiencia exitosa – central virtual de aprendizaje.

Eje temático: smartcampus y tendencias tecnológicas emergentes en la universidad (aplicaciones de la inteligencia artificial, blockchain, analítica, IoT, impresión 3D, experiencia inmersiva).

1. Introducción

Telesalud - Universidad de Caldas y su área de educación continua han construido un modelo de educación virtual replicable que aporta a la automatización de sus procesos, por tal motivo se desarrolló un módulo basado en inteligencia artificial, integrado a la central virtual de aprendizaje, el cual ha sido entrenado a partir del análisis de los datos de registro de notas y comportamiento de los estudiantes en cada cohorte de las ofertas académicas existentes. Dicho software se centra en predecir el riesgo de fracaso o deserción por bajos promedios de calificación (inspirados en la producción cinematográfica "sentencia previa" del año 2002 - dirigida por Steven Spielberg quien utiliza un oráculo para predecir el riesgo de las personas en cometer un crimen)^[9], tomando medidas inmediatas según el comportamiento de los estudiantes virtuales, y el grupo objetivo en su mayoría son profesionales de la salud, profesionales de carreras afines al área de la salud, estudiantes de últimos semestres de carreras enfocadas a la salud y agentes comunitarios, y que por su naturaleza de aporte a la salud de un colectivo, requieren la certificación de sus habilidades enfocadas al servicio de la comunidad. Gracias a las funcionalidades de la máquina se logra mayor dedicación en tiempo y recursos, enfocados en la calidad de los contenidos académicos, acompañamiento, motivación y efectividad del modelo educativo.

Los resultados del programa de Educación Virtual de la Universidad de Caldas ha sido postulado en diferentes eventos y premios a nivel nacional e internacional y Telesalud ha sido catalogado como un grupo con altos estándares de calidad y compromiso dentro de la comunidad de profesionales de la salud del país, es así, como esta experiencia en este modelo ha permitido obtener diferentes reconocimientos nacionales e internacionales, como el premio Latinoamericano Frida/eLAC (2010, 2013 y 2015), por los logros obtenidos con el programa de Telemedicina y Educación Virtual. En el 2012 el BID, la CEPAL, la OPS/OMS entre otros, otorgaron el reconocimiento al grupo con las "Mejores Prácticas en Telesalud". En el 2014, Telesalud recibió el reconocimiento a la Innovación en Salud Pública en Latinoamérica entregado por parte de eSAC por la oferta del Curso Virtual en Salud Infantil, destacándose los bajos índices de deserción (proyecto regional financiado por el International Development Research Centre (IDRC) - diseñado e implementado conjuntamente por la OPS/OMS y el Grupo de Investigación en Salud, Equidad e Innovación de la Universidad de Toronto, Canadá). Reconocimiento por el impacto de Educación virtual en Colombia y Latinoamérica por RedClara y Tical (2018) y ganadores de los PREMIOS CÓNDOR (2018) que entrega la Red de Emprendimiento Tecnológico en el departamento de Caldas. Dentro de los principales resultados globales en educación virtual se encuentran los beneficios obtenidos a través del uso de diferentes herramientas virtuales educativas como lo son: clases en vivo streaming, participación en eventos virtuales, aplicativos de gamificación, capacitación y/o certificación para más de 50.000 profesionales de la salud en más de 800 municipios de Colombia y 10 países de Latinoamérica, además de 20 ofertas virtuales permanentes en cursos virtuales, seminarios, simposios y diplomados en diversos temas de alta relevancia para el aporte social, convirtiendo a Telesalud en un instrumento de beneficios en salud, para la comunidad en Colombia y Latinoamérica.

⁹ K. Dick, P. (2002). *Minority Report* [Film]. Hollywood: Steven Spielberg.

La implementación de un modelo de educación virtual predictiva es la evolución tecnológica a la administración manual y compleja de todo un proceso basado en datos, que dificultaba al operador integrar la parte humanística y emotiva al acompañamiento académico de los estudiantes, ahora con el modelo predictivo el equipo puede preocuparse en incrementar la calidad de los componentes académicos, la motivación 1 vs 1, el acompañamiento permanente, y la analítica de datos es una prioridad para la mejora continua. Es por ello, que la ventaja de los participantes es *"el aprendizaje automático que pretende aportar a los programas de Inteligencia Artificial, la capacidad de adaptarse sin tener que ser reprogramados"*^[10], y así permitir tiempo para considerar el conocimiento y la motivación en las ofertas académicas, con un excelente resultado no solamente en los promedios de calificación sino también en las competencias adquiridas, en el impacto sobre la atención de los pacientes, la humanización en los servicios de atención, la respuesta de la comunidad, esto gracias a la calidad de los recursos de aprendizaje y certificación de habilidades.

¿Cómo se desarrolló el modelo? Telesalud - Universidad de Caldas diseñó y desarrolló el modelo de educación predictiva por medio de un software de IA que es adaptable y reprogramable, todo ello con el objetivo primordial de centrar los esfuerzos en el conocimiento, este desarrollo permitirá obtener mejores resultados en los estudiantes, centrar la mirada en las falencias académicas y mejorar o alivianar los factores externos ajenos a la academia, enfocando al participante como el actor más importante en el proceso académico.

Problemática y contexto de propuesta del proyecto

Pensar constantemente en los patrones de comportamiento de un participante virtual se ha convertido en la razón de ser del programa de Telesalud, independientemente de desarrollar una central virtual de aprendizaje, donde el participante encuentre de manera intuitiva los objetos virtuales de aprendizaje, y el proceso de obtención de aprendizaje sea fácil e integral, aunque existen connotaciones diversas que presenta una población que requiere de una academia diferente y una andragogía totalmente adaptada a las necesidades propias de un sector que presenta variables de deserción como falta de tiempo, la familia, el trabajo y otros proyectos, sumados a *"los problemas de 'deslaborización', acoso laboral, pérdida de la autonomía, exceso de carga laboral, contrataciones ilegales, baja remuneración y desempleo"*^[11], además de otras variables que afectan directamente los procesos de educación, encontrando que si no se implementa un modelo predictivo basado en algoritmos inteligentes, continuará siendo muy complejo sortear las dificultades que presentan los modelos virtuales tradicionales, como lo son:

- Dificultades para identificar el perfil y comportamiento de cada estudiante.
- Dificultades de soporte técnico permanente.
- Dificultades de soporte académico permanente.

¹⁰ CAMPO NIEVES, L. (2011). Aplicación de la inteligencia artificial en la predicción de fracasos endodónticos. protocolo de investigación [Ebook] (1st ed.). Madrid.

¹¹ Empleo, E. (2019). Panorama laboral de los médicos en Colombia. Retrieved from <https://www.empleo.com/co/noticias/investigacion-laboral/panorama-laboral-de-los-medicos-en-colombia-5835>

- Problemas para la identificación de las causas de abandono.
- Dificultad para tomar medidas correctivas antes que ocurra el abandono.
- Dificultades de comunicación efectiva durante la capacitación.
- Dificultades para obtener una respuesta positiva y oportuna de los participantes a los comunicados de acompañamiento durante la capacitación.
- Dificultades de acompañamiento y motivación oportuna.
- Cobertura total de las eventualidades reportadas por los participantes.
- Capacidad de respuesta por parte del recurso humano para despejar las solicitudes de soporte a grandes contingentes.
- Dificultad para identificar oportunamente bajos promedios que generen riesgo de certificación.
- Dificultades para activar los protocolos de recuperación, habilitación y rescate.
- Dificultad para analizar datos a través de protocolos de operación manual.
- Dificultades de caracterización para evitar la deserción.
- Dificultades para predecir los resultados finales de la capacitación virtual.
- Dificultades en la implementación del plan de acción y mapas de riesgo.
- Dificultades para automatizar un mayor número de tareas.
- Dificultades para medir la capacidad operativa del recurso humano.

También se encontraron eventualidades basadas en los aspectos culturales y dificultades personales, que se puede definir en primera medida en uno de los cuellos de botella que presenta la operación del programa, es por ello, que la solución presenta ventajas que ayudan en gran medida a predecir en tiempos determinados el comportamiento y los promedios de las calificaciones de los participantes para permitir el desarrollo esperado de su capacitación sin que los elementos distractores le impidan llevar a cabo el buen curso del aprendizaje y logrando contar con tiempos flexibles y suficientes para brindar soluciones oportunas de respuesta a los compromisos académicos.

Para el Ministerio de Salud es muy importante *“La utilización de las competencias profesionales, requeridas para el desempeño efectivo y con calidad de los profesionales de la salud, como parte de los referentes básicos para su formación, lo cual exige cambios y ajustes importantes en los procesos formativos”*^[12], es por esto, que para Telesalud Universidad de Caldas y su modelo de Educación Virtual es muy importante la anterior premisa y creen firmemente que centrarse en la calidad de la educación y las facultades intrínsecas y extrínsecas del estudiante, es el camino para formar, ayudar, motivar, e impartir el conocimiento integral a los participantes y así salvar vidas a millones de pacientes en Colombia y Latinoamérica.

Es así, como la central virtual de aprendizaje integra una herramienta tecnológica que a través de analítica de datos, basada en la lógica y la inteligencia artificial predice el alcance en promedio de calificaciones y establece el protocolo según patrones de comportamiento y perfiles definidos, con la apropiación de algoritmos que inician el proceso alerta y recuperación del estudiante por medio de SMS, llamadas telefónicas y

¹² GAVIRIA URIBE, A., CORREA SERNA, L., DÁVILA GUERRERO, C., BURGOS BERNAL, G., ORTÍZ MONSALVE, L., & BARRERA GUAUQUE, O. et al. (2016). Perfiles y competencia profesionales en salud [Ebook] (1st ed.). Bogotá: MinSalud. Retrieved from <http://www.minsalud.gov.co>

correos electrónicos, completando un ciclo que permite equilibrar los niveles de aprendizaje esperados del participante en los módulos correspondientes de cada oferta académica. Con este Modelo Predictivo de Educación Virtual de Telesalud Universidad de Caldas se asegura que el participante siempre se encuentre dentro de los promedios que garanticen la certificación, de la mano con el acceso oportuno al conocimiento y flexibilidad en las opciones de aprendizaje, asegurando en un 95% que los factores externos como descuido, dificultades, tiempo o actitud no interfieran negativamente, gracias a la acción oportuna del sistema predictivo que activa el protocolo de acompañamiento para el rescate de los estudiantes en riesgo.

2. Descripción de la solución tecnológica implementada

El modelo de Educación virtual tiene diferentes componentes que integran la Central Virtual de Aprendizaje, y en ella se desarrolla la estrategia que implementada va encaminada al desarrollo integral del participante de las capacitaciones virtuales, se parte de la credibilidad en las personas y el conocimiento de las dificultades en el contexto laboral y las necesidades de capacitación que el medio exige, además del conocimiento de primera mano de las especificaciones de los perfiles categorizados, que brindan a Telesalud y a su software predictivo valiosa información que permite predecir y adelantar a la culminación positiva previo a protocolos de rescate, si así lo requiere cada participante: *"De esta premisa viene el esfuerzo de Telesalud y su sistema integral de educación virtual, de este virtuoso sentir nos integra y nos fusiona con las necesidades de cada profesional de la salud y a llevar el mensaje a través de las manos de una enfermera, el análisis de un especialista o el concepto médico, todo ello traducido en esperanza a alguien o al país que con cada clic hacemos crecer"*^[13].

Para ello se cuenta en primera instancia con el desarrollo de la Central Virtual de Aprendizaje (CVA) que integra diferentes soluciones tecnológicas como el software de Inteligencia Artificial que permite predecir el riesgo de fracaso por bajos promedios de calificación evolucionando en la gestión de datos, realizando protocolos de semaforización y de algoritmos de alarma y rescate para el estudiante, ilustrados en la siguiente imagen:

¹³ Granada Aguirre, A., Carmona Gallego, C., & Parra Alzate, H. (2018). Impacto del programa de Educación Virtual en Telesalud Universidad de Caldas (p. 62). Cartagena de Indias: María José López Pourailly, RedCLARA.



Fig. 1. Rutas de implementación del modelo predictivo.

El módulo de educación virtual predictiva tiene el siguiente funcionamiento: cuando un alumno ha culminado su tercer módulo de estudio, éste módulo indica con un 75% de acierto en qué clase terminará el estudiante, y se activan los protocolos automatizados del área de soporte para realizar el acompañamiento respectivo a los estudiantes; a medida que el alumno avanza en sus módulos, el porcentaje de acierto aumenta, hasta llegar a un máximo de 84.45 % una vez haya culminado su quinto módulo. Hay que aclarar que para fines de experimentación se realizaron clasificaciones con los 6 módulos completos y se alcanzó un acierto del 99%.



Fig. 2. Proceso de implementación de educación virtual predictiva. Aquí se presenta el proceso necesario para la implementación del módulo de educación virtual predictiva, integrado a la Central Virtual de Aprendizaje.

A continuación, se describen cada uno de los pasos ejecutados en el proceso de implementación del módulo de educación virtual predictiva:

- **Selección de información:** En este paso se escogió los registros de 611 alumnos de diferentes cursos que tuvieran diferentes resultados de desempeño final.
- **Revisión de la base de conocimiento y limpieza de los datos:** Luego de obtener una base de conocimiento, se revisa la calidad de los datos y se realiza la tarea de limpiar los datos anómalos o nulos, también se agrupan los datos para que cada columna del modelo final represente un módulo ejecutado.
- **Creación de las clases:** Después de ajustar el modelo, se analiza con el equipo de trabajo cuales son los umbrales de calificación para establecer las clases que serán predecidas, y se crean 3 clases las cuales son rojo que corresponde a alumnos que no se certifican o se certifican con una calificación baja, amarillo son los alumnos que obtienen una nota medio alta, y verde son los alumnos que obtienen una calificación alta.
- **Selección de algoritmos:** Para la ejecución de este proyecto se escogieron inicialmente los algoritmos ID3, Naive bayes, Perseptrón multicapa, y regresión logística.
- **Experimentación de los algoritmos con la base de conocimiento:** El entrenamiento y validación de cada algoritmo se realizó de la siguiente manera: se utilizó el método de validación cruzada con 10 folders usando

el total de los registros (611), se realizaron varias iteraciones para conocer cual es el número de módulos mínimo que debió completar un estudiante y así predecir en que clase terminaría al final de la capacitación, es decir, primero 6 módulos, luego 5 módulos y así hasta 2 módulos.

- **Integración del algoritmo de mejor desempeño con la CVA:** Luego de comparar los porcentajes de acierto de cada uno de los algoritmos se decide implementar el algoritmo de regresión logística, que en un mínimo de 3 módulos completados se puede predecir en un 75.94% en que clase puede terminar un alumno. En la sección de resultados obtenidos e impactos se mostrará el desempeño de los demás algoritmos. Hay que aclarar que entre más módulos complete el alumno, el algoritmo tiene un porcentaje de acierto más alto, para la implementación del algoritmo se utilizó el lenguaje de programación R y su gran cantidad de librerías que facilitaron la implementación.

3. Aspectos críticos y relevantes a resaltar y detallar

Los aspectos críticos de la solución que se deben resaltar se encuentran relacionados con la implementación y la puesta en marcha, como bien se sabe para poder realizar un nuevo desarrollo se requiere de una inversión de recursos económicos y destinación de recurso humano para tal fin, así como la probabilidad de no lograr obtener los resultados esperados, inicialmente para el entrenamiento de una máquina es indispensable contar con datos suficientes para poder dar paso a la analítica de los mismos y soportados en los resultados de los ejercicios de prueba en entornos simulados para lograr contar con resultados positivos para la nueva solución, se pasó de una solución netamente manual a una automatizada que evidencia las dificultades de los estudiantes semaforizados y categorizados, con estadísticas y probabilidades de falla o no certificación, implementando estrategias que permiten humanizar la educación con más tiempo para desarrollar labores de carácter académico y mantenimiento de plataforma.

Se resalta específicamente desde la parte técnica que "hay muchos casos en los que se pueden resolver situaciones complejas haciendo uso de reglas deterministas, hasta el punto que su uso consigue sistemas automáticos que se comportan como humanos expertos en un dominio particular, permitiendo tomar decisiones, por ejemplo: en sistemas de control de tráfico, transacciones bancarias, o diagnóstico de enfermedades. Entre las opciones disponibles, los sistemas basados en reglas se han convertido en una de las herramientas para tratar de manera eficiente una buena colección de problemas, ya que las reglas deterministas constituyen la más sencilla de las metodologías utilizadas en sistemas expertos. En estos sistemas, la base de conocimiento de la que se parte contiene las variables y el conjunto de reglas que definen el problema, y el motor de inferencia son capaces de extraer conclusiones aplicando métodos de la lógica clásica sobre esta base. Una regla en este contexto es una proposición lógica que relaciona dos o más objetos del dominio e incluye dos partes, la premisa y la conclusión, que se suele escribir normalmente como "Si premisa, entonces conclusión". Cada una

de estas partes es una expresión lógica con una o más afirmaciones objeto - valor conectadas mediante operadores lógicos (y, o, o no)^[14].

4. Resultados obtenidos y su impacto

Tabla 1. Valores de la calificación media que obtiene cada una de las clases en un módulo determinado.

Módulos	Rojo	Amarillo	Verde
Módulo 1	10	11,74	13,36
Módulo 2	9,17	12,46	14,18
Módulo 3	9,95	12,8	14,19
Módulo 4	11,86	13,76	14,52
Módulo 5	8,84	11,72	13,96
Módulo 6	8,1	12,35	13,66

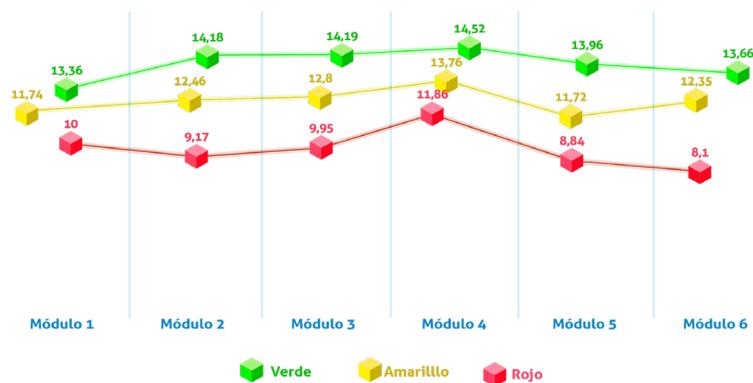


Fig. 3. Calificación media obtenida por cada clase en un módulo determinado. La gráfica anterior nos permite apreciar de mejor manera cual es el comportamiento que tienen los alumnos de cada una de las clases.

¹⁴ Caparrini, F., Caparrini, F., & Work, W. (2019). Sistemas Basados en Reglas - Fernando Sancho Caparrini. Retrieved from <http://www.cs.us.es/~fsancho/?e=103>

Tabla 2. Porcentaje de aciertos de cada algoritmo con un número mínimo de módulos terminados.

Etiquetas de fila	2 Módulos	3 Módulos	4 Módulos	5 Módulos	6 Módulos
ID3	65,95	71,68	71,35	78,39	86,09
Naive Bayes	66,44	69,88	73,48	78,72	84,78
Perceptrón multicapa	66,45	74,96	77,25	82,48	99,02
Regresión Logística	68,41	75,94	78,55	84,45	99,67
Promedio Aciertos	66,8125	73,115	75,1575	81,01	92,39

La anterior tabla, muestra el desempeño de todos los algoritmos con cada uno de los módulos, como primer criterio para escoger el algoritmo a implementar se tiene en cuenta que el porcentaje de acierto esté por encima de la media de desempeño, en ese caso los que cumplen con porcentajes muy similares son el perceptrón multicapa y la regresión logística en todos los módulos, y el segundo criterio fue la cantidad de recursos necesarios para ejecutar el algoritmo y es ahí donde presenta un menor consumo de recursos el algoritmo de regresión logística. Por lo tanto, el algoritmo que cumple mejor con los criterios es la regresión logística.

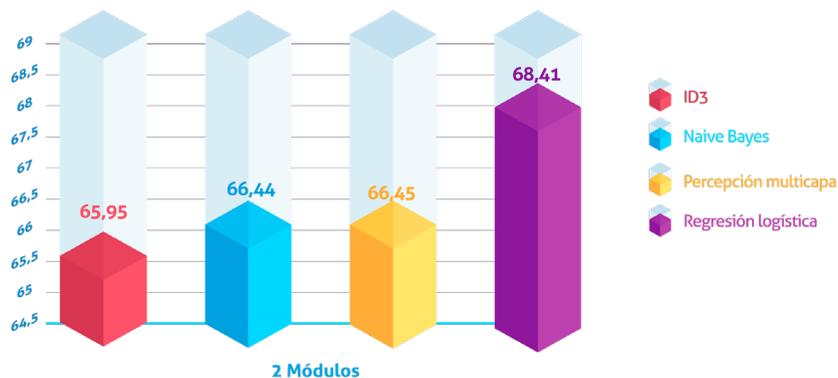


Fig. 4. Porcentaje de cierto de los algoritmos con 2 módulos. Respecto el porcentaje de acierto presentado en esta gráfica se establece que se presenta mucha incertidumbre al momento de realizar una predicción.

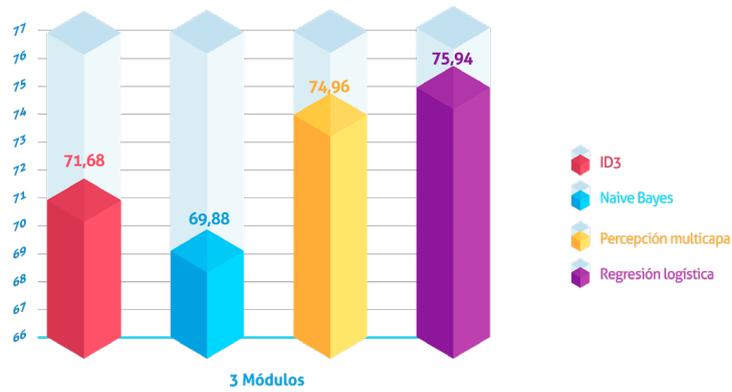


Fig. 5. Porcentaje de acierto de los algoritmos con 3 módulos. Como se mencionó anteriormente el algoritmo seleccionado es la regresión logística y se acepta que el error de clasificación presente en una etapa temprana del desarrollo del curso sea cerca del 25 %.

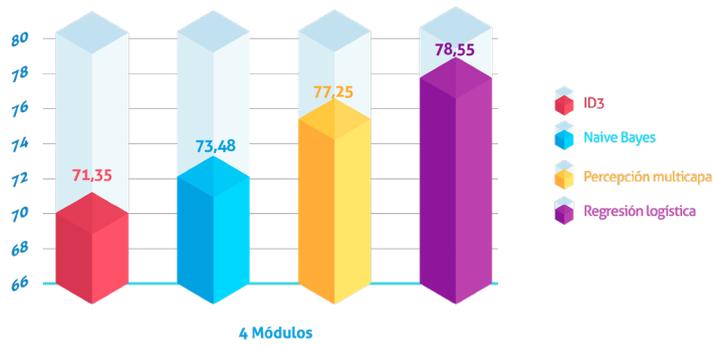


Fig. 6. Porcentaje de acierto de los algoritmos con 4 módulos.



Fig. 6. Porcentaje de acierto de los algoritmos con 5 módulos.

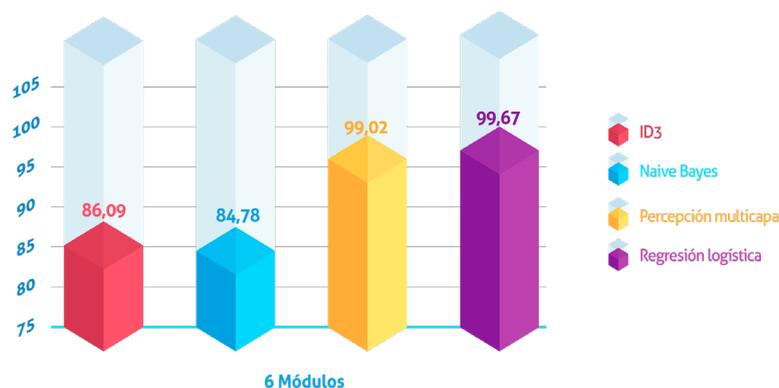


Fig. 7. Porcentaje de acierto de los algoritmos con 6 módulos.

Las gráficas 5, 6 y 7 muestran que los algoritmos perceptrón multicapa y regresión logística son los que muestran un mejor desempeño en el porcentaje de acierto, también se puede apreciar que a medida que los alumnos completan sus módulos, los algoritmos tienen un mejor porcentaje de acierto llegando a tener aciertos por encima del 99 %.

5. Aprendizajes

Inicialmente los registros académicos se realizaban de manera manual, los documentos de Excel eran la herramienta por predilección y el que por medio de colores se catalogaba los estudiantes y generaban los porcentajes correspondientes, este método generaba datos matemáticos exactos, pero hoy en día la educación virtual predictiva permite de manera anticipada el perfil de los participantes de los cursos, simposios y diplomados que se imparten y con base en esa información un software inteligente basado en procesos matemáticos se aprendió realmente que el tiempo invertido en formatos de Excel podía ser sustituido por un sistema organizado y predictivo que sumara sus horas al cálculo anticipado del comportamiento de perfiles determinados, retroalimentando los conceptos en cada oferta que se realiza en Telesalud Universidad de Caldas, además se aprendió que las notas y actividades deben ser siempre enfocadas en los perfiles se conocen y se evidencian en el software, entendiendo que cada una de las actividades se diseña no por el grado de complejidad sino que por la analítica de datos que logra diseñar una actividad enfocada en el porcentaje de conocimiento que brinda al estudiante, reevaluando la práctica de evaluaciones difíciles, ya que el software brinda el aprendizaje y el concepto de la individualidad del estudiante, lo categoriza y lo encasilla en un perfil que genera la formación de manera óptima con flexibilidad, teniendo en cuenta la motivación constante hacia el objetivo que es la certificación de las capacitaciones, entonces pues la reflexión a esta labor desde la Universidad de Caldas "es que las instituciones no se queden en la etapa de lo conocido, busquen la innovación para mejorar, obtener datos que permitan afianzar la relación con el estudiante, bajar el nivel de la brecha que supone un programa de educación

virtual, humanizando desde los datos que nos proveen los software y sobre todo buscar sistemas que automaticen los procesos humanizando además la educación, el sistema predictivo produce información que permite tener tiempo de clasificar y tratar al estudiante como un individuo con características diferentes y no como un grupo generalizado.

6. Impactos no esperados

Al tener perfiles definidos ya evidenciados por los datos obtenidos de la herramienta del modelo predictivo a través de la aplicación de encuestas, se encontraron diferentes hallazgos que no tenidos en cuenta y que son inherentes al factor humano al que se quiere llegar, el ser humano no tiene la perfección de un software predictivo y presenta márgenes de error que a la hora de implementar modelos matemáticos que deben ser tenidos en cuenta:

- Una de las variables no esperadas mostró que un 35% de los estudiantes tuvieron dependencia al sistema y a los recordatorios de sus actividades, esto evidenció un perfil diferente que brinda datos para catalogarlos e incluirlos en un nuevo perfil que requiere mayor acompañamiento para lograr la certificación.
- Se evidenció en un 5% que como se establecieron metas mínimas, los participantes adaptaron su comportamiento a sólo obtener las notas mínimas que les ayudara a lograr la certificación y otro grupo de estudiantes (5%) cayeron sistemáticamente en la escala de riesgo, al generar la costumbre de esperar la alarma para ingresar de nuevo y lograr pasar con promedio aceptable para certificación.

Entre las desventajas principales de la herramienta que han llevado a complementarlo o sustituirlo con otros procedimientos de razonamiento se destacan:

- Relaciones opacas entre reglas: aunque las reglas de producción son muy simples desde un punto de vista individual, las interacciones que se producen a larga distancia entre la red de reglas existentes pueden ser muy opacas, lo que hace que generalmente sea difícil saber qué papel juega una regla en particular en la estrategia global de razonamiento que hay detrás.
- Estrategias de búsqueda muy ineficientes: esencialmente, el motor de inferencia realiza una búsqueda exhaustiva en todas las reglas en cada ciclo de iteración, por lo que los sistemas de reglas con muchas reglas (que pueden llegar a ser miles) son lentos y, a menudo, inviables en problemas del mundo real^[15].

¹⁵ Caparrini, F., Caparrini, F., & Work, W. (2019). Sistemas Basados en Reglas - Fernando Sancho Caparrini. Retrieved from <http://www.cs.us.es/~fsancho/?e=103>

Agradecimientos

La Universidad de Caldas, la Facultad de Ciencias para la Salud y su Decana Dra. Claudia Patricia Jaramillo, ya que apoyó de manera incondicional este y muchas otras iniciativas de educación virtual y desarrollo de software de inteligencia artificial que logra el libre desarrollo, la evolución e innovación en los procesos en pro de la comunidad universitaria y el país.

Los autores desean de manera cordial y emotiva agradecer al sr. Germán González Martínez por su participación en la recopilación de los datos estadísticos y revisión del presente documento y al sr. Andrés Felipe Aristizábal por la elaboración de gráficos estadísticos que permite hacer de éste documento agradable para el lector, además al equipo de ingeniería de Telesalud Universidad de Caldas por su increíble desarrollo que permite un modelo predictivo adaptado al modelo de educación virtual, y a todo el equipo interdisciplinar de trabajo por la construcción diaria de un proyecto innovador que salva vidas a través de la tecnología en Colombia y Latinoamérica.

Agradecemos a Red RENATA Red Nacional Académica de Tecnología Avanzada por la infraestructura de red de tecnología avanzada y conexión de alta velocidad que permite desde las universidades el desarrollo de proyectos colaborativos de ciencia, educación e innovación al igual que una conexión estable y segura para los participantes virtuales en el continente latinoamericano, y a la RedCLARA por realizar convocatorias que incentivan y hacen posible que las personas que integran las universidades construyan e innoven en beneficio de la comunidad académica iberoamericana.

Referencias

1. K. Dick, P. (2002). *Minority Report* [Film]. Hollywood: Steven Spielberg.
2. CAMPO NIEVES, L. (2011). *Aplicación de la inteligencia artificial en la predicción de fracasos endodónticos. protocolo de investigación* [Ebook] (1st ed.). Madrid.
3. Empleo, E. (2019). Panorama laboral de los médicos en Colombia. Retrieved from <https://www.empleo.com/co/noticias/investigacion-laboral/panorama-laboral-de-los-medicos-en-colombia-5835>
4. GAVIRIA URIBE, A., CORREA SERNA, L., DÁVILA GUERRERO, C., BURGOS BERNAL, G., ORTÍZ MONSALVE, L., & BARRERA GUAUQUE, O. et al. (2016). *Perfiles y competencia profesionales en salud* [Ebook] (1st ed.). Bogotá: MinSalud. Retrieved from <http://www.minsalud.gov.co>
5. Granada Aguirre, A., Carmona Gallego, C., & Parra Alzate, H. (2018). *Impacto del programa de Educación Virtual en Telesalud Universidad de Caldas* (p. 62). Cartagena de Indias: María José López Pourailly, RedCLARA.
6. Caparrini, F., Caparrini, F., & Work, W. (2019). *Sistemas Basados en Reglas - Fernando Sancho Caparrini*. Retrieved from <http://www.cs.us.es/~fsancho/?e=103>

UnaCloud: Ejecutando Aplicaciones HPC en las Salas de Computadores de la Universidad

Carlos E. Gómez^{a,b}, Jaime Chavarriga^a, Harold Castro^a

^a Departamento de Ingeniería de Sistemas y Computación
Universidad de los Andes
Bogotá, Colombia
{ce.gomez10, ja.chavarriga908, hcastro}@uniandes.edu.co

^b Universidad del Quindío
Armenia, Colombia

Resumen. La Computación de Alto Rendimiento (HPC) ofrece grandes capacidades de cómputo para una diversidad de aplicaciones científicas. Lamentablemente, la infraestructura requerida para este tipo de computación es muy costosa. En lugar de usar equipos especializados, UnaCloud permite ejecutar estas aplicaciones en los computadores de las salas y laboratorios de la universidad, aprovechando los recursos disponibles. En promedio, el uso de CPU en estos computadores no supera el 5%. UnaCloud ofrece el modelo de Infraestructura como Servicio (IaaS) soportado exclusivamente en la capacidad restante de estos computadores para ejecutar máquinas virtuales y crear clústeres virtuales bajo demanda que pueden ejecutar aplicaciones HPC. A diferencia de sistemas similares, UnaCloud busca ofrecer un servicio que va más allá del simple mejor esfuerzo para permitir la posibilidad de ejecutar de manera desatendida aplicaciones de tipo bolsa de tareas, e incluso habilitar la ejecución exitosa de aplicaciones paralelas que utilizan librerías como MPI. En la actualidad usamos este sistema como una alternativa de bajo costo cuando la infraestructura HPC no tiene disponibilidad o se requiere ejecutar aplicaciones no tan exigentes. Este artículo describe nuestra experiencia, mostrando detalles de su implementación, los logros obtenidos y algunas lecciones aprendidas en los últimos años.

Eje temático: Tecnologías en la innovación educativa e investigación.

1. Introducción, Problemática y Contexto

Universidades de todo el mundo usan esquemas de Computación de Alto Rendimiento (High Performance Computing, HPC) para realizar simulaciones y procesamiento complejo soportando investigaciones en temas tan variados como economía, análisis genético, química computacional y física computacional.

Lamentablemente, la infraestructura usada para ejecutar aplicaciones HPC es muy costosa. Además de los costos de hardware, que suelen ser superiores a los 150 mil dólares, es necesario adecuar espacios físicos y conseguir personal especializado para administrar la infraestructura. Muchas universidades no pueden costear este tipo de sistemas o solo pueden adquirir sistemas muy pequeños.

En contraste, casi todas las universidades cuentan con salas y laboratorios con computadores que se usan en clases y que, durante grandes períodos de tiempo, permanecen subutilizados y con altas capacidades ociosas [2] [4] [7]. Por ejemplo, un estudio de 2011 en la Universidad de los Andes mostró que el uso de procesador en dos salas de computadores oscilaba entre 5% y 8% durante las horas del día y entre el 3% y el 5% en horas de la noche [7] (ver Fig. 1). Un análisis similar en 2018 mostró un consumo de CPU entre 4% y 6% en el día, y cerca del 2% en la noche [6] (ver Fig. 2). Una evaluación reciente de algunos de estos computadores (Intel Core i7-7700 @3.60Hz) muestran una capacidad de 47 GigaFLOPS para una máquina virtual usando cuatro núcleos de procesamiento de los ocho disponibles en cada computador. Al consolidar la capacidad ociosa de los computadores de las salas donde se ejecuta UnaCloud es posible alcanzar a un máximo de 5.5 TeraFlops, similar a la capacidad de los equipos de HPC más pequeños de hace varios años.

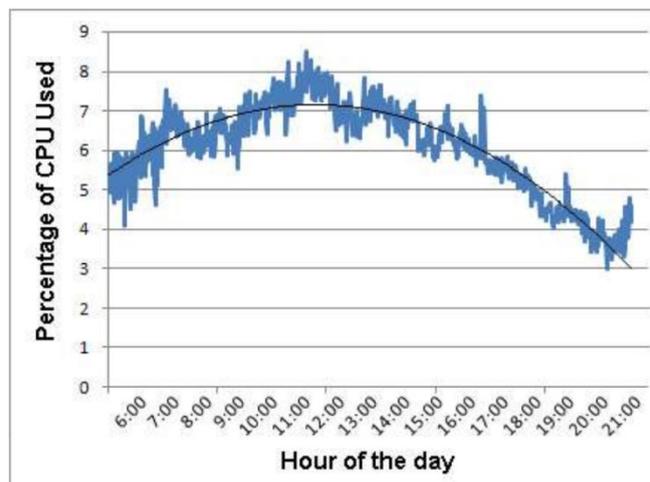


Fig. 1. Consumo promedio de CPU 7:00 am – 10:00 pm, 2011 [7]

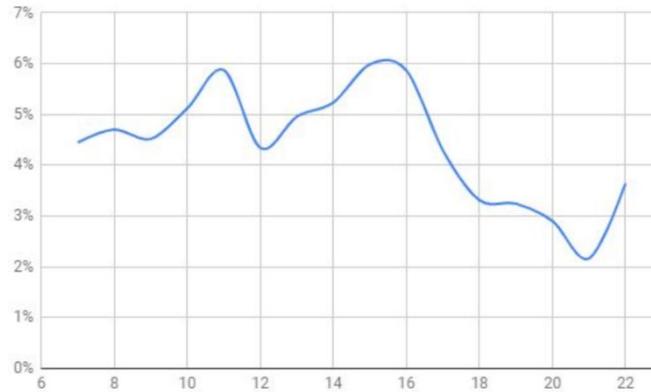


Fig. 2. Consumo promedio de CPU 7:00 am – 10:00 pm, 2018 [6]

UnaCloud [9] [10] es un sistema oportunista que ofrece servicios de cloud computing aprovechando los recursos disponibles en los computadores en estas salas y laboratorios de computadores. Utilizando máquinas virtuales que se ejecutan en baja prioridad en esos escritorios, permiten a los investigadores y académicos crear clústeres virtuales y ejecutar simulaciones y aplicaciones usando tecnologías de HPC, entre otros. Aunque no permite lograr el mismo poder de las máquinas HPC más poderosas, UnaCloud permite ejecutar experimentos en computadores de escritorio que de otra forma serían desaprovechados.

Este artículo describe nuestra experiencia usando UnaCloud para ejecutar aplicaciones HPC. Aquí presentamos detalles de su implementación en la Universidad de los Andes, los logros obtenidos y algunas lecciones aprendidas en los últimos años.

En el resto del artículo está organizado como se enumera a continuación. La sección 2 describe UnaCloud y la ejecución de las aplicaciones HPC. La sección 3 muestra aspectos críticos y relevantes. La sección 4 expone los resultados obtenidos y su impacto. La sección 5 habla de aprendizajes, aciertos y errores. La sección 6 presenta los impactos no esperados. Finalmente, la Sección 7 presenta conclusiones y trabajo futuro.

2. Descripción de la Solución Tecnológica Implementada

UnaCloud es una plataforma desarrollada al interior de la Universidad de los Andes para ofrecer Servicios de Infraestructura (IaaS) usando la capacidad ociosa de los equipos en salas y laboratorios de la universidad [9] [10]. Esta sección describe esta plataforma y su uso para la ejecución de aplicaciones HPC.

Plataforma UnaCloud

Desde 2010, el grupo de investigación COMIT ha estado trabajando en el desarrollo de UnaCloud, una plataforma oportunista que ofrece servicios de computación en la nube

sobre computadores de escritorio (opportunistic desktop cloud). Esta plataforma ejecuta máquinas virtuales (MVs) sobre computadores de escritorio, incluso cuando esos computadores son usados por estudiantes y empleados. Básicamente, UnaCloud es una plataforma oportunista ya que busca aprovechar los recursos de computación disponibles en esos computadores sin que los usuarios noten que se están ejecutando trabajos adicionales. Para hacer esto posible, la plataforma ejecuta las máquinas virtuales usando la más baja prioridad posible.

UnaCloud utiliza una arquitectura cliente-servidor. Por un lado, la plataforma ofrece una aplicación web donde los usuarios pueden solicitar la ejecución de clústeres y máquinas virtuales. Las peticiones son administradas por un servidor que mantiene un inventario de todos los equipos de escritorio disponibles, asigna las máquinas virtuales a esos equipos y controla la ejecución de estas máquinas virtuales. Por otro lado, hay un software agente de UnaCloud en cada uno de los computadores de escritorio. Este agente recibe instrucciones del servidor, descarga copias de los discos de máquinas virtuales (a.k.a., imágenes virtuales), inicia, detiene y genera copias de respaldo de máquinas virtuales. La Fig. 3 da una visión general de UnaCloud, y la Fig. 4 presenta la arquitectura del sistema.

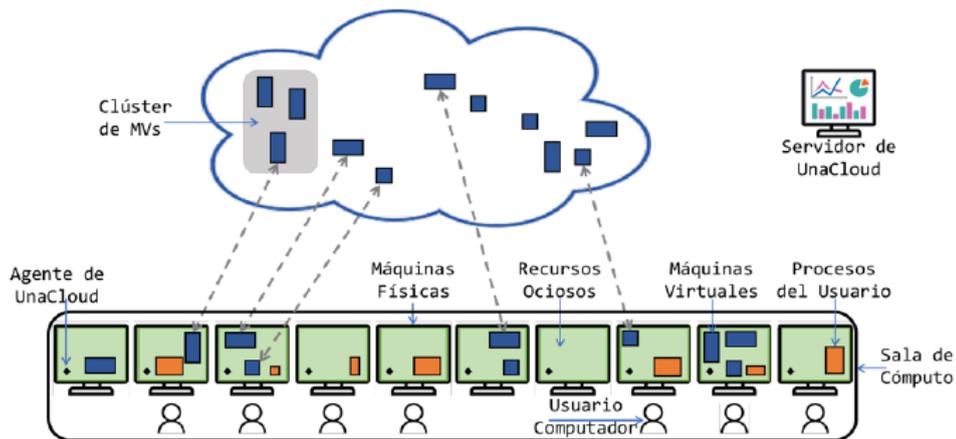


Fig. 3. Visión general de UnaCloud.

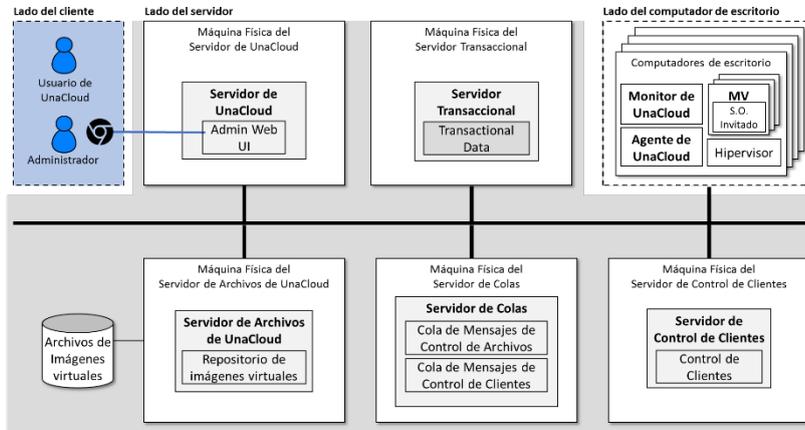


Fig. 4. Arquitectura de UnaCloud

Uno de los principales problemas de este tipo de plataformas es que son muy sensibles a la intervención de los usuarios trabajando en los computadores. Por ejemplo, estos usuarios pueden iniciar aplicaciones intensivas en cómputo que afecten la ejecución de las máquinas virtuales. Además, pueden desconectar la red o reiniciar el computador. Todo esto puede causar interrupciones en el procesamiento y, en algunas ocasiones, perder el trabajo realizado hasta el momento. Por esto mismo, un aspecto clave ha sido el desarrollo reciente de mecanismos para mejorar la confiabilidad de UnaCloud. En los últimos años se han desarrollado opciones para generar copias (snapshots) de las máquinas virtuales para poder reiniciar o migrar el procesamiento a otros equipos de escritorio de ser necesario [3].

El desarrollo de UnaCloud es realizado por miembros del grupo COMIT: investigadores post-doctorales y estudiantes de doctorado, maestría y pregrado. A partir de los diferentes requerimientos y avances tecnológicos, se definen y realizan proyectos que extienden la plataforma y corrigen errores y problemas en el funcionamiento. En la actualidad, por ejemplo, se encuentran en desarrollo proyectos para integrar nuevas formas de tolerancia a fallos, el uso de contenedores docker, además de máquinas virtuales, para la ejecución de aplicaciones científicas y la posibilidad de integrar a la capacidad de cómputo ofrecida las tarjetas de GPUs disponibles en una de las salas. El software tiene una licencia de código abierto y está disponible en Github¹⁶.

Plataforma UnaCloud en la Universidad de los Andes

En la Universidad de los Andes, UnaCloud está desplegado en dos salas de computadores, la sala Waira y la sala Turing, utilizadas por estudiantes de la Escuela de Sistemas y Computación. Las Tablas 1 y 2 muestran las especificaciones de los computadores en esas salas.

¹⁶ <https://github.com/UnaCloud/UnaCloud2>

Tabla 1. Especificaciones de los computadores de la Sala Waira.

Característica	Especificación
Cantidad de computadores	78
Procesador	Intel (R) Core (TM) i7-7700 CPU @ 3.60Hz [8 core(s) x64]
Memoria RAM	32 GB RAM DDR4 @1.20 GHz
Disco duro	1 TB
Tarjeta de video	Intel HD Graphics 630
Tarjeta de red	Intel 1 Gb/s
Sistema operativo	Microsoft Windows 10 Enterprise

Tabla 2. Especificaciones de los computadores de la Sala Alan Turing

Característica	Especificación
Cantidad de computadores	39
Procesador	Intel (R) Core (TM) i7-7700 CPU @ 4.20Hz [8 core(s) x64]
Memoria RAM	64 GB RAM DDR4 @1.20 GHz
Disco duro	1 TB
Tarjeta de video	Nvidia Quadro P600 Graphics
Tarjeta de red	Intel 1 Gb/s
Sistema operativo	Microsoft Windows 10 Enterprise

Todos los computadores de estas salas tienen instalado el agente de UnaCloud y el hipervisor Oracle VirtualBox 5.2.x. Este software es utilizado por UnaCloud para ejecutar las máquinas virtuales.

En nuestra infraestructura, los usuarios de UnaCloud pueden crear sus propias imágenes de máquina virtual. Por ejemplo, un usuario puede usar VirtualBox en su computador personal para crear una máquina virtual instalando un sistema operativo y algunas aplicaciones. Luego, la imagen del disco duro puede subirse a la plataforma y ser usada para crear clústeres con máquinas virtuales en las salas de computadores. En la actualidad, hemos hecho despliegues exitosos de clústeres de hasta 448 máquinas virtuales en esos equipos de escritorio.

UnaCloud en Operación

A continuación, presentamos algunas imágenes de la plataforma en ejecución.

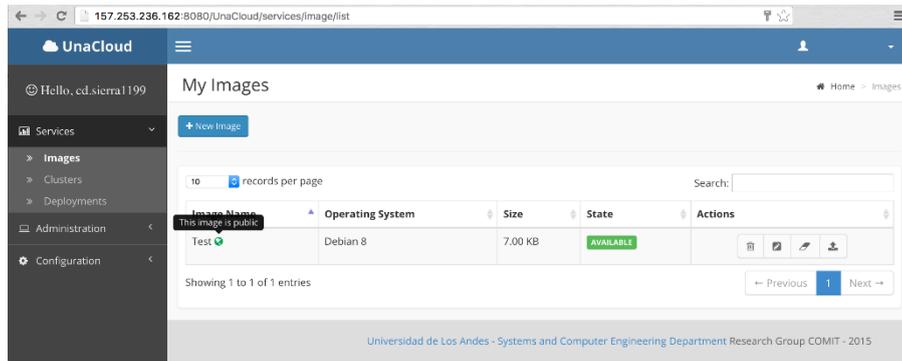


Fig. 5. Administración de imágenes virtuales en UnaCloud

La Fig. 5 muestra a la izquierda el menú de opciones. Si se trata de un usuario básico, el menú ofrece los servicios básicos de operación: administración de imágenes virtuales, clústeres y despliegues. En cambio, si el usuario es administrador (como el de la figura), el menú incluye funciones de administración y configuración. En la parte derecha de la Fig. 5 se puede ver el panel de administración de imágenes virtuales.

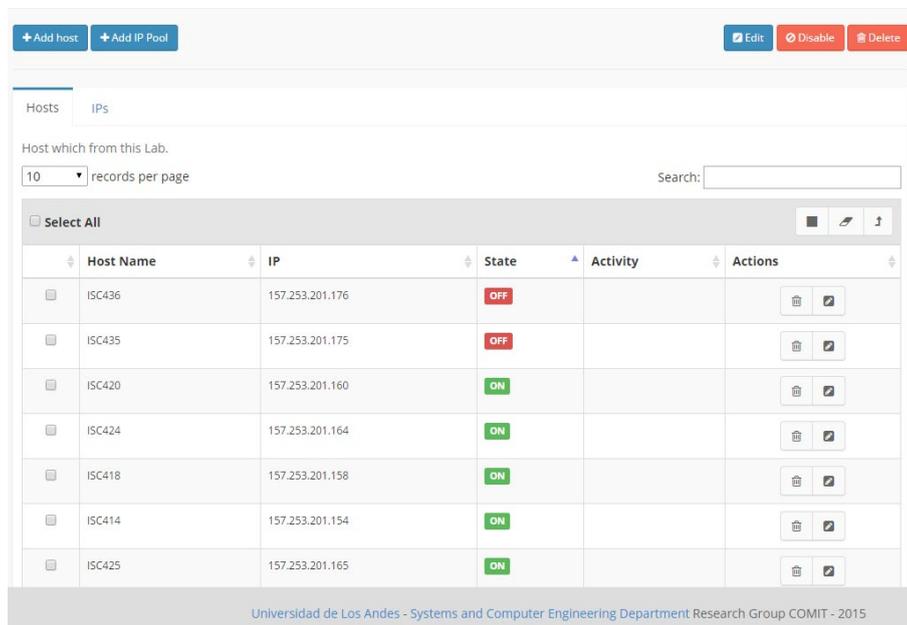


Fig. 6. Estado de las máquinas físicas en UnaCloud.

La Fig. 6 muestra el estado de las máquinas físicas de una sala en UnaCloud. Se puede ver un listado con los nombres de las máquinas físicas, las direcciones IP de las máquinas físicas, el estado (ON/OFF), y los botones de eliminar y editar. En la parte superior hay unos controles para adicionar máquinas físicas a la sala y direcciones IP al pool de direcciones que se pueden asignar a las MVs.

Ejecución de Aplicaciones HPC en UnaCloud

Tradicionalmente estas infraestructuras son utilizadas para ejecutar aplicaciones desacopladas en forma de bolsas de tareas (Bag-of-Tasks). Si bien UnaCloud ha sido utilizado exitosamente con este tipo de trabajos, en los últimos años hemos querido ofrecer un servicio más avanzado y los investigadores han estado usando aplicaciones HPC como GROMACS y algunos programas personalizados que están basadas en librerías MPI (Message Passing Interface). Estas aplicaciones pueden ejecutarse en un clúster de computadores sin memoria compartida. En esos casos, el software corre al mismo tiempo en todos los nodos intercambiando información usando protocolos de red. Además, se usan discos compartidos en red para que todos los nodos accedan a los mismos archivos para leer datos de entrada y almacenar los resultados.

Una ejecución típica de una aplicación HPC en UnaCloud implica el uso de un servidor NFS para compartir archivos. En este servidor, se crea una carpeta con los programas y los datos de entrada que serán usados por todos los equipos del clúster. En cada equipo se tiene instalado el software a usar (p.ej. GROMACS), software de comunicación SSH y las librerías de MPI. Todos los nodos son configurados para que cuenten con un usuario determinado (p.ej. unacloud) y puedan conectarse usando llaves asimétricas entre ellos sin necesidad de ingresar la contraseña. Este usuario es usado para ejecutar el programa correspondiente. Al iniciar el programa, de forma automática, este se conecta a los otros computadores, inicia el programa en todos los nodos y hace la ejecución de las tareas intercambiando datos a través de la red.

Para ejecutar aplicaciones HPC hemos creado imágenes de máquinas virtuales con los sistemas operativos Linux Debian y Ubuntu Linux, incluyendo software NFS, MPI (OpenMPI y MPICH) y otras aplicaciones científicas. El usuario puede crear un clúster usando estas imágenes. Se puede usar un servidor NFS en una de esas máquinas virtuales o en otro computador de la universidad. La obtención final de los resultados depende del tipo de programa y de la forma como se usa el disco compartido en NFS.

En este tipo de aplicaciones paralelas, si uno de los nodos falla, todo el proceso se interrumpe. No importa si los otros nodos siguen funcionando. Esto es de gran relevancia en UnaCloud debido a que los equipos pueden verse afectados por las acciones de los usuarios en los computadores de escritorio. Si el usuario en el computador desconecta el computador o reinicia alguno de los computadores, el procesamiento debe ser reiniciado. Aplicaciones como GROMACS ofrecen mecanismos de checkpointing que permiten reanudar las tareas sin perder todo el trabajo previo. Además, en UnaCloud se han implementado varias técnicas, como checkpointing a nivel de MPI [1] y snapshot global [3] a nivel de máquinas virtuales, para permitir reanudar tareas en aplicaciones que no proveen checkpointing de forma nativa.

3. Aspectos Críticos y Relevantes

UnaCloud ha estado en operación durante casi 10 años. Durante este tiempo hemos identificado algunos aspectos claves que nos han permitido usar la plataforma con éxito.

Bajo Uso de la Capacidad de Cómputo. Durante todo este tiempo, diversos estudios han mostrado que los computadores en los laboratorios y salas tienen, en promedio, un bajo uso de sus capacidades de cómputo [1] [2] [6] [7]. En su gran mayoría, los usuarios usan aplicaciones como Microsoft Office, navegadores web, y software de edición de diagramas y documentos, que no consumen toda la capacidad de procesamiento de los computadores actuales. Algunos usuarios usan lenguajes de programación que consumen casi toda la capacidad durante procesos de compilación, pero en tiempos muy cortos. Existen pocas excepciones, como usuarios que utilizan software de simulación, optimización matemática o de machine learning. Soluciones como UnaCloud son mucho más efectivas en salas de computadores con capacidad ociosa. Una universidad podría establecer salas de computadores para aplicaciones con alto consumo de CPU, donde no se desplegarían máquinas virtuales, y salas destinadas a aplicaciones menos exigentes donde sí se podrían crear clústeres virtuales.

Uso de Computadores de Gran Capacidad en las Salas. En la actualidad es posible contar con salas con computadores económicos, con bajas capacidades, que funcionan como terminales o escritorios remotos de otros servidores. En estos equipos, aunque se tengan bajos niveles de utilización de CPU, la capacidad ociosa puede ser insuficiente para ejecutar máquinas virtuales. En la Universidad de los Andes las salas utilizadas para ejecutar las máquinas virtuales cuentan con computadores de alta capacidad. Son computadores con procesadores Intel Core i7 (con 8 núcleos virtuales de procesamiento) y 32 GB de RAM donde más de una máquina virtual de 1 a 4 procesadores virtuales y con 1 a 8 GB de RAM pueden ejecutarse sin afectar de manera significativa a los otros programas corriendo en el mismo equipo.

Actualización Permanente de la Infraestructura. Para poder ejecutar apropiadamente las aplicaciones HPC se necesita, no solo computadores de gran capacidad, sino también sistemas de red que funcionen apropiadamente. El uso de switches de gran capacidad en las diferentes salas permite lograr comunicaciones de alto rendimiento entre los nodos, beneficiando el funcionamiento de las aplicaciones HPC. Además, la actualización permanente de los equipos reduce la probabilidad de fallo de hardware como discos duros y unidades de alimentación. En la Universidad de los Andes, la infraestructura de las salas usadas en UnaCloud, incluyendo computadores y sistemas de red, han sido actualizadas tres veces en los últimos nueve años.

Actualización Permanente de la Plataforma. Constantemente aparecen nuevas tecnologías de virtualización y de manejo de clústeres y redes virtuales. En muchos casos, estos permiten aprovechar mejor el hardware disponible y lograr menores niveles de interferencia sobre los trabajos que realizan los estudiantes en las salas. Por esta razón, UnaCloud es una plataforma que se ha estado actualizando permanentemente.

En la actualidad, por ejemplo, se está explorando el uso de hipervisores tipo 1 en sistemas operativos de escritorio, como HyperV en Windows 10 y Hypervisor Framework en OSX, y el uso de contenedores Docker. El software implementado abstrae el concepto de hypervisor lo que habilita la integración de nuevos sistemas fácilmente.

Administración de la Plataforma. Finalmente, otro aspecto crítico es la administración de la plataforma. En el caso de la Universidad de los Andes, gran parte del trabajo se realiza con asistentes graduados y estudiantes. Lamentablemente esto causa que se tenga una alta rotación de personal: los estudiantes trabajan solo por uno o dos semestres y en algunos casos se tienen períodos de tiempo en donde se cuenta con muy pocas personas para apoyar el desarrollo y la operación de la plataforma.

Nótese que uno de los beneficios de una plataforma como UnaCloud es poder contar con capacidad de cómputo HPC sin necesidad de invertir en infraestructura especializada. Aunque este modelo puede requerir una mayor inversión en las salas de cómputo, para tener equipos y redes de alta capacidad, esta inversión permite ejecutar aplicaciones de alto rendimiento usando la misma infraestructura, a un menor costo.

4. Resultados Obtenidos y su Impacto

Durante todo este tiempo, UnaCloud ha permitido la realización de una gran variedad de proyectos de investigación. Por un lado, ha permitido investigaciones de simulación en temas como ingeniería civil, biología computacional, química computacional y procesamiento de imágenes. Por otro lado, ha permitido desarrollar una variedad de proyectos de investigación y desarrollo en ingeniería de sistemas y computación.

Proyectos que Han Usado UnaCloud

Existe una gran variedad de proyectos de investigación que han usado UnaCloud en los últimos años. Los siguientes son algunos ejemplos:

En Ingeniería Civil, UnaCloud ha sido usado para evaluar algoritmos escalables para la planeación y nivelación de recursos [8]. En este proyecto, se evaluaron los algoritmos usando una infraestructura distribuida combinando máquinas virtuales en una plataforma tradicional de cloud computing (ISCloud basada en VmWare ESX) y máquinas virtuales en UnaCloud.

En Procesamiento de Imágenes, UnaCloud fue utilizada en el proyecto KMC ANTs Template para realizar un número significativo de pruebas, ajustar los parámetros de sus algoritmos y realizar los cálculos finales sin usar los computadores del laboratorio especializado Colivri ni afectar los procesos que se llevan habitualmente en ese lugar.

En Ingeniería Química, UnaCloud fue usado en el proyecto `\textit{BSGrid}` donde se aplicaron técnicas de química computacional para evaluar el impacto de varios biopesticidas sobre la bacteria *Bacillus thuringiensis* [5].

En Bioinformática, UnaCloud ha sido integrado al software LONI pipeline para ejecutar de forma coordinada múltiples herramientas de procesamiento bioinformático. Esta plataforma ha sido usada para crear clústeres virtuales y realizar procesamiento en varios proyectos.

Proyectos de Investigación Sobre UnaCloud

Además de estos proyectos, al interior de los programas de Ingeniería de Sistemas y Computación, Maestría en Ingeniería de Sistemas y el Doctorado en Ingeniería, se han realizado varios proyectos de investigación y desarrollo sobre la plataforma UnaCloud. Entre estos proyectos, se han realizado:

Ocho proyectos de grado enfocados en temas como la automatización del copiado y despliegue de máquinas virtuales, la optimización de la transferencia de imágenes virtuales, el perfilamiento de consumo de recursos en salas de cómputo, y la utilización de contenedores Docker.

Siete proyectos de maestría enfocados en extensiones a UnaCloud para soportar mejor las aplicaciones HPC, soluciones para el análisis de secuencias genéticas, la interoperabilidad con otras plataformas de cloud computing y el soporte para la virtualización de GPUs,

Dos tesis doctorales, una enfocada en la asignación de recursos de cómputo considerando la eficiencia energética y otra más en el mejoramiento de la confiabilidad de la plataforma.

Como resultado, tenemos más 10 artículos en memorias de conferencias, 2 artículos de revistas y 2 capítulos de libros que muestran la investigación desarrollada en UnaCloud.

5. Aprendizajes – Aciertos y Errores

Son muchas las lecciones aprendidas en todo este tiempo. A continuación, citamos algunas de las más relevantes.

Los Recursos Mejoran Día a Día. Cada vez que se renuevan los equipos, los nuevos llegan con mejores especificaciones que los anteriores y se pueden aprovechar mejor. Por ejemplo, constantemente aparecen nuevos procesadores y arquitecturas de computadores más rápidos y con mayor capacidad de procesamiento. Actualizar la infraestructura nos permite contar con mejores capacidades para la ejecución de máquinas virtuales.

La Red y el Almacenamiento Son Cuellos de Botella. Mientras las infraestructuras especializadas en HPC usan sistemas de comunicación de alta velocidad, memoria compartida y discos de alto rendimiento, la plataforma UnaCloud usa infraestructura de menor capacidad. Cualquier mejora que se pueda hacer en las redes usadas en las salas de computadores o en los discos en esos computadores es bastante beneficiosa para la ejecución de aplicaciones en plataformas oportunistas como UnaCloud.

El Monitoreo es Complejo pero Esencial. Debido a la naturaleza no dedicada de los computadores de las salas de cómputo, estos equipos fallan con frecuencia debido a la intervención de los usuarios. Es posible que las máquinas virtuales se interrumpan una o dos veces en esos computadores. Es importante contar con un sistema de monitoreo que permita detectar los fallos y tomar decisiones que permitan continuar con la ejecución de las tareas. Además, información histórica de los fallos podría ayudar a generar indicadores y planificar mejor el aprovisionamiento de las máquinas virtuales. Lamentablemente, durante todos estos años, lograr un sistema apropiado de monitoreo

ha sido elusivo. Ha sido difícil no solo determinar las variables que son apropiadas para tomar decisiones, sino capturar esos valores, almacenar los datos y procesar las métricas correspondientes para obtener información útil que nos ayude a mejorar la confiabilidad del sistema. Este es un tema de investigación abierto.

Es Importante Contar con Ambientes de Pruebas. Los desarrolladores no pueden depurar apropiadamente el software de UnaCloud en las salas de cómputo debido a que estas salas de cómputo son administradas por personal diferente. Por ejemplo, por motivos de seguridad, no es posible revisar algunos mensajes de auditoría, analizar los permisos de las aplicaciones o depurar las máquinas virtuales en ejecución. Para poder hacer el desarrollo y pruebas de nuevas versiones, en la actualidad estamos usando un laboratorio de redes donde podemos configurar, analizar y revisar los equipos, las redes y las máquinas virtuales.

Integrar los Proyectos de Investigación no es Fácil. Debido a que el desarrollo de UnaCloud se realiza mediante proyectos de grado y de investigación, la integración de las mejoras y cambios en el código fuente no siempre es fácil. Muchos de estos proyectos terminan con un artículo de investigación y una prueba de concepto de software. En realidad, estos cambios no se incluyen en la rama principal de software y requieren de actividades adicionales de desarrollo y pruebas. Es recomendable contar con un equipo de arquitectos y desarrolladores de software.

6. Impactos no Esperados

Entre los impactos no esperados, existen dos que quisiéramos resaltar. Por un lado, algunas tareas de UnaCloud han tenido un impacto negativo en la infraestructura de red de la Universidad. El diseño general de la red mantiene los servidores de UnaCloud en un extremo y las salas de computadores en el otro. La transmisión de imágenes virtuales, el almacenamiento de datos NFS y algunas tareas de copias de respaldo requieren pasar por una gran cantidad de switches y equipos de red. Cuando se hacen tareas complejas de procesamiento, el tráfico asociado puede afectar segmentos de red en otras áreas de la Universidad. En el último año hemos estado trabajando en unos ajustes a la topología de red de la Universidad que nos permita tener los servidores de UnaCloud con una conexión directa a las salas de computadores.

Por otro lado, cuando se inició el proyecto de UnaCloud, no consideramos que la presencia del usuario en los mismos computadores causara tantos inconvenientes en la ejecución de las máquinas virtuales. Uno de los impactos no esperados es que el desarrollo de la plataforma nos ha permitido realizar varias investigaciones sobre la detección de fallos, el perfilamiento de hardware acorde con estos fallos y la implementación de diferentes técnicas de tolerancia a fallos y mejoramiento de la confiabilidad. Hemos debido implementar y ajustar técnicas que se aplican en data-centers y soluciones privadas de cloud computing en software de virtualización usados en computadores de escritorio como Oracle VirtualBox.

7. Conclusiones y Trabajo Futuro

UnaCloud es una plataforma que ofrece servicios de cloud computing aprovechando la capacidad ociosa en computadores de escritorio en salas y laboratorios. Usando esta plataforma, un usuario puede crear clústeres y máquinas virtuales que se ejecutan en esos computadores, incluso aunque estos equipos están siendo usados por estudiantes o trabajadores en su quehacer diario. En los últimos años, esta plataforma ha sido usada para realizar una gran variedad de experimentos científicos, usando esquemas de computación de alto rendimiento (HPC), sin necesidad de utilizar la infraestructura HPC de la Universidad que es más costosa y muchas veces está ocupada en otros proyectos. UnaCloud ha sido utilizado con éxito en varios proyectos en áreas como ingeniería civil, ingeniería química, procesamiento de imágenes y bioinformática.

Entre los elementos claves para uso de la plataforma hemos identificado que la actualización permanente de los equipos y la infraestructura de la sala nos ha permitido contar con mayor capacidad de cómputo para los proyectos. Además, la actualización de la plataforma nos ha permitido aprovechar mejor los avances en las tecnologías de virtualización y contar con personal de apoyo ha sido clave para la realización de los proyectos.

Durante este tiempo, hemos estado trabajando en el monitoreo del sistema para detectar fallos y poner en práctica estrategias de tolerancia a fallos. Este es un dolor de cabeza permanente en el desarrollo de la plataforma. Hemos planeado trabajo futuro en la implementación de un nuevo modelo de monitoreo que permita determinar con mayor seguridad las causas de las fallas, en nuevos modelos de tolerancia a fallos y en el uso de contenedores en lugar de esquemas tradicionales de máquinas virtuales. Además, estamos muy interesados que otras universidades utilicen la plataforma e iniciar proyectos conjuntos con otras entidades para aprovechar la experiencia que tenemos ejecutando aplicaciones HPC en computadores de escritorio.

Agradecimientos

Queremos agradecer la participación en el proyecto UnaCloud a quienes durante estos años han participado con sus ideas y aportes. Particularmente, queremos hacer un reconocimiento especial a Jesse Padilla, David Bonilla, Aurelio Vivas, Eduardo Rosales, Germán Sotelo, César Forero, César Díaz y Mario Villamizar.

Referencias

1. E. Bohorquez, E. Rosales, and H. Castro. Running MPI applications over an opportunistic infrastructure. In *2015 Ninth International Conference on Complex, Intelligent, and Software Intensive Systems*, pages 446–453. IEEE, 2015.
2. P. Domingues, P. Marques, and L. Silva. Resource usage of windows computer laboratories. In *Parallel Processing, 2005.ICPP 2005 Workshops. International Conference Workshops on*, pages 469–476. IEEE, 2005.
3. C. E. Gomez, H. E. Castro, and C. A. Varela. Global snapshot of a distributed

system running on virtual machines. In *2017 29th International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD)*, pages 169–176. IEEE, 2017.

4. C. E. Gomez, C. O. Diaz, C. A. Forero, E. Rosales, and H. Castro. Determining the real capacity of a desktop cloud. In *Latin American High Performance Computing Conference*, pages 62–72. Springer, 2015.

5. A. Gonzalez, H. Castro, M. Villamizar, N. Cuervo, G. Lozano, S. Restrepo, and S. Orduz. Mesoscale modeling of the bacillus thuringiensis sporulation network based on stochastic kinetics and its application for in silico scale-down. In *2009 International Workshop on High Performance Computational Systems Biology*, pages 3–12. IEEE, 2009.

6. N. Gomez and J. G. Tamura. Detección y predicción de fallas a partir de métricas de performance en sistemas desktop cloud. Tesis de maestría. Universidad de los Andes, 2018.

7. A. A. Oviedo. UnaCloud MSA: Plataforma basada en UnaCloud para la generacion y análisis de alineamientos múltiples de secuencias. Tesis de maestría, Universidad de los Andes, 2011.

8. J. L. Ponz-Tienda, A. Salcedo-Bernal, and E. Pellicer. A parallel branch and bound algorithm for the resource leveling problem with minimal lags. *Computer-Aided Civil and Infrastructure Engineering*, 32(6):474–498, 2017.

9. E. Rosales, H. Castro, and M. Villamizar. UnaCloud: opportunistic cloud computing infrastructure as a service. *Cloud Computing*, pages 187–194, 2011.

10. UnaCloud. Unacloud: Opportunistic cloud computing platform, 2018. Accessed 20-04-2018.

Implementación de un Sistema de Gestión de la Investigación (CRIS) utilizando tecnologías abiertas: El caso de la Universidad Autónoma de San Luis Potosí

Rosalina Vázquez Tapia ^a,

^a Dirección de Biblioteca Virtual Universitaria - Sistema de Bibliotecas, Universidad Autónoma de San Luis Potosí,

Niño Artillero S/N Zona Universitaria CP 78290, San Luis Potosí, S.L.P., México
alinavn@uaslp.mx, alinante@gmail.com

Resumen. En este trabajo se presenta primeramente un marco conceptual sobre el origen, definición y características de un Sistema de Gestión de la Investigación o CRIS (Current Research Information System) y su importancia como una estrategia sostenida para el poblamiento de un Repositorio Institucional de Acceso Abierto. Se plantea el modelo de Sistema CRIS diseñado para la Universidad Autónoma de San Luis Potosí, México, y el desarrollo del proyecto por etapas. Se describe la arquitectura del Sistema y las tecnologías abiertas implementadas para el desarrollo de cada uno de sus componentes técnicos. Finalmente, se señalan los resultados obtenidos hasta el momento y algunas reflexiones sobre los retos y aprendizajes obtenidos.

Abstract. In this paper a conceptual framework is first presented about the origin, definition and features of a Current Research Information System (CRIS) and its relevance as a sustainable practice for populating an Open Access Institutional Repository. The CRIS System model is proposed and designed for the Autonomous University of San Luis Potosí, Mexico, as well as the project development in stages. The System architecture is described and the open technologies implemented for each one of the technical components. Finally, the obtained results so far are noted, as well as some observations about the challenges and learned lessons.

Palabras Clave: Institutional Repository, Current Research Information System, ORCID.

Eje temático: Tecnologías en la innovación educativa e investigación.

1. Introducción

En la última década, como una consecuencia del movimiento de Acceso Abierto (*Open Access*) establecido en el 2003 a través de la Declaración de Berlín ¹⁷, se ha expandido y fortalecido el concepto de Ciencia Abierta (*Open Science*) o Ciencia Digital (*Digital Science*) definida por la Comisión Europea como "una radical transformación de la naturaleza de la ciencia e innovación debido a la integración de las ICT al proceso de investigación y a una cultura de internet abierta y de colaboración" [1]. La Ciencia Abierta no sólo contempla a las dos vías del Acceso Abierto como medios para la publicación y diseminación de los resultados de investigación, sino que también incluye el uso de recursos abiertos, software abierto, métricas alternativas, web social, datos abiertos y licencias abiertas, por mencionar algunos, en cada una de las etapas del ciclo de vida de la investigación.

Como parte de este paradigma de investigación abierta, surgen los llamados CRIS (*Current Research Information System*) o RIMS (*Research Information Management System*) como un nuevo modelo para organizar y analizar, desde una perspectiva integradora, toda la información relativa a la actividad científica [2].

Un RIM es la integración, curación y utilización de información acerca de la investigación y está emergiendo como un área de creciente interés y relevancia en muchas bibliotecas universitarias [3]. Por su parte, un CRIS es un sistema de información que almacena y gestiona datos de las actividades de investigación en una institución. Es aquella herramienta que permite gestionar de manera global todos los procesos relacionados con la investigación; facilita al investigador la realización de su CV y genera informes de gestión, informes a los financiadores, bibliografía de investigación, etc. [4]

El propósito de un CRIS es proveer a los investigadores, gestores, administradores, financiadores y tomadores de decisiones, información confiable y sólida acerca de las actividades y resultados de investigación para el desarrollo de estrategias institucionales. Los sistemas CRIS son también útiles para evaluar el desarrollo de la investigación, exponer los resultados para encontrar nuevas oportunidades de financiamiento, evitar la duplicidad de actividades, analizar tendencias, referenciar al texto completo o publicaciones académicas multimedia, localizar nuevos contactos e identificar nuevos mercados para productos de investigación y diseminar la investigación para promover el conocimiento abierto [5].

Por lo general, un CRIS almacena y gestiona información en torno a cuatro entidades principales: investigadores, organizaciones, proyectos y resultados de investigación. Los tipos de datos se muestran en la Fig. 1.

¹⁷ Declaración de Berlín sobre Acceso Abierto al Conocimiento en Ciencias y Humanidades. <http://www.mpg.de/en>



RVT, Septiembre de 2018

Fig. 1. Tipos de datos de un CRIS. Fuente: Elaboración propia.

Para describir de manera normalizada cada uno de estos datos (metadatos) y sus atributos, así como hacerlos interoperables, EuroCRIS¹⁸, una asociación internacional sin fines de lucro que reúne a expertos en investigación y sistemas CRIS, creó el formato CERIF (*the Common European Research Information Format*), un modelo estándar de datos de investigación recomendado por la Unión Europea a sus Estados miembros para facilitar la interoperabilidad de los sistemas CRIS.

El formato CERIF es un modelo de dominio completo que permite una descripción formal de muchos aspectos inherentes al dominio de investigación. Define un formato en XML para el intercambio de datos (*CERIF XML*) que ha sido adaptado por OpenAIRE como la regla o guía para la recolección (*harvesting*) e importación de metadatos de sistemas CRIS. En las guías de interoperabilidad para sistemas CRIS v. 1.0 de OpenAIRE [6] se define el esquema específico de metadatos en CERIF XML que comprende ocho entidades: Publication; Product/Dataset; Person; Organisation; Project; Funding; Equipment; Service.

Para implementar un CRIS, además de utilizar estándares de metadatos, se requiere la utilización de vocabularios controlados (diccionarios de términos) para la descripción (sintaxis) de los metadatos y el uso de identificadores digitales persistentes para la identificación correcta de los autores, sus trabajos académicos y su producción científica. Ejemplos de estos identificadores son el DOI para artículos, ISBN para libros, ISSN para revistas, Ringgold para organizaciones y ORCID (*Open Research and Contributor Identifier*)¹⁹ para autores.

Por otro lado, otro componente importante de la Ciencia Abierta que posibilita la disseminación de la producción científica y académica de una institución u organización

¹⁸ EuroCRIS. <https://www.eurocris.org/>

¹⁹ ORCID. <https://orcid.org/>

son los Repositorios Institucionales, que constituyen la vía verde del Acceso Abierto (AA). Un Repositorio Institucional (RI) es un conjunto de servicios que la universidad ofrece a los miembros de su comunidad para la gestión y para la disseminación de los materiales en forma digital creados por la institución y sus miembros [7].

En este sentido, de acuerdo a COAR [8] una de las mejores prácticas para poblar un Repositorio Institucional es la creación e interconexión con sistemas CRIS, que ha resultado para los investigadores un valor agregado o incentivo (*"la zanahoria"*) para contribuir al Repositorio y que junto con la implementación de un Mandato Institucional²⁰ (*"el garrote"*), constituyen una estrategia eficaz para la sostenibilidad de los repositorios institucionales.

Referente a la infraestructura requerida para la implementación de un CRIS y su posible interoperabilidad con un Repositorio Institucional, se deben considerar por una parte, la instalación de un software en un servidor de producción o máquina virtual (solución en la nube), y por otra, la implementación de un protocolo o software adicional para conectar el CRIS con el Repositorio Institucional. En cuanto a las soluciones de software, pueden ser de cuatro tipos: a) Comercial (Ej.: Pure de Elsevier, Converis de Clarivate Analytics); b) Open source (Ej.: Dspace-CRIS, VIVO); c) Desarrollo propio; d) Híbrido.

Para conectar un CRIS con un Repositorio Institucional existen varias alternativas. Si ya se cuenta con un plataforma de software para la gestión del repositorio como lo es DSpace, lo más viable es configurar el protocolo SWORD que permite el depósito del documento y sus metadatos en el repositorio. Para implementar un repositorio con las funcionalidades de un CRIS fue creado el software open source Dspace-CRIS, que es una extensión de DSpace para la gestión y visibilidad de datos y documentos de investigación.

1.1 Justificación y contexto del proyecto

De acuerdo al marco teórico planteado anteriormente, resulta claro que la implementación de un CRIS vinculado a un Repositorio Institucional, representa para una Institución de investigación o educación superior y sus miembros, importantes ventajas, beneficios y oportunidades tales como:

- Incrementar el número de citas y por ende el factor de impacto de las publicaciones
- Obtener métricas alternativas de uso en la web
- Generar informes e indicadores de producción científica a nivel institucional
- Incrementar el impacto y la visibilidad académica de la investigación
- Contar con una plataforma institucional para almacenar, gestionar y preservar los productos de investigación y las hojas de vida de sus investigadores.
- Construir una fuente información confiable para los procesos de evaluación de la investigación, promoción de personal académico y toma de decisiones.
- Promover el auto archivo y poblamiento sostenido del Repositorio Institucional.

²⁰ Un Mandato Institucional es una política de AA emitida por una Institución que establece el depósito obligado de las publicaciones y productos de investigación producidos por sus miembros en el Repositorio Institucional.

Con el propósito de construir una estrategia institucional de Acceso Abierto y brindar en el mediano y largo plazo estos beneficios a la comunidad académica de la Universidad Autónoma de San Luis Potosí (UASLP)²¹, por iniciativa de la Biblioteca Virtual Universitaria (BVU) del Sistema de Bibliotecas, fue creado en el 2012 el Repositorio Institucional NINIVE²² de la UASLP. El diseño de NINIVE comprende cinco comunidades en función de los tipos de acervos producidos: Archivo Institucional, Documentación Científica, Recursos Educativos, Patrimonio Documental, Publicaciones y Producciones Universitarias. El software de gestión sobre el cual está soportado es DSpace V. 6.2.

Actualmente (mayo de 2019), en el Repositorio Institucional NINIVE están disponibles en texto completo un total de 2,230 recursos digitales, de los cuales se encuentran 1719 tesis de licenciatura y posgrado, 228 libros publicados por la Editorial Universitaria, 198 artículos de revistas de divulgación y algunos recursos educativos y documentos de fondo antiguo.

En lo que respecta al contexto institucional, la UASLP es una Institución Pública de Educación Superior que cuenta con una matrícula de 31,102 estudiantes y una planta académica de 3,340 profesores de los cuales 807 son de tiempo completo y 485 pertenecen al Sistema Nacional de Investigadores. La oferta educativa está integrada por 100 carreras de licenciatura y 90 de posgrado de los cuales 69 están reconocidos en el Padrón Nacional de Posgrados de Calidad del Consejo Nacional de Ciencia y Tecnología (CONACYT). En cuanto a la producción científica de los últimos 5 años, se tiene identificados un total de 8,650 productos de investigación científica publicados y arbitrados y 317 productos de investigación publicados por la Institución bajo el esquema de licenciamiento de Acceso Abierto.

De esta producción científica, sólo un 15% aproximadamente se encuentra depositado en NINIVE. Esto puede deberse a los siguientes factores:

- La política de auto archivo del Repositorio Institucional NINIVE no ha tenido el impacto esperado debido por una parte, a la falta de políticas institucionales que lo promuevan y faciliten, y por otra, a que los investigadores lo consideran como un trabajo adicional sin beneficios claros en el corto plazo.
- La producción científica se encuentra dispersa y muy poco visible. No se cuenta con ningún sistema o base de datos institucional para la gestión de la investigación que permita a los investigadores administrar su hoja de vida y a la vez, darle visibilidad a sus publicaciones.
- El proceso de búsqueda y recuperación de las publicaciones de los investigadores en las bases de datos es manual, lo que implica varios meses de trabajo para la generación de indicadores.
- No hay un identificador único para el investigador salvo su clave de empleado que para efectos de interoperabilidad no es válida. Tampoco existen catálogos de autores, revistas, proyectos o fuentes de financiamiento, relacionados a las actividades de investigación.

²¹ Universidad Autónoma de San Luis Potosí. <http://www.uaslp.mx>

²² Repositorio Institucional NINIVE. <http://ninive.uaslp.mx>

- Hay un desconocimiento general entre los investigadores sobre las políticas editoriales de las revistas comerciales en las que publican y por ende sobre los derechos patrimoniales que ceden o podrían retener.

Para atender la problemática anterior, crear servicios de valor agregado para los investigadores orientados a la ciencia abierta y lograr la sostenibilidad operativa y financiera del Repositorio Institucional NINIVE en el largo plazo, a finales del 2015 se formuló un proyecto institucional con tres objetivos estratégicos:

1. Elaborar una propuesta de Mandato Institucional de Acceso Abierto, para el depósito en NINIVE de una copia digital de las publicaciones y producciones universitarias.
2. Desarrollar un Sistema de Gestión de la Investigación (CRIS) como parte del modelo de servicios del Repositorio Institucional NINIVE.
3. Diseñar un plan de sensibilización y capacitación dirigido principalmente a los investigadores, así como también, implementar estrategias de difusión y cooperación a nivel nacional e internacional.

Para el diseño e implementación del Sistema de Gestión de la Investigación, primeramente se llevó a cabo un estudio de *benchmarking* que consistió en recopilar, analizar y evaluar información en fuentes electrónicas de información, para identificar el tipo de tecnologías e infraestructura requerida y las mejores prácticas y modelos de referencia de sistemas CRIS en instituciones educativas de Estados Unidos y España. Con base en los resultados de este estudio y el análisis previo de las necesidades y características propias de la UASLP en materia de investigación, se diseñó un modelo propio (Véase Fig. 2.) basado en los siguientes principios de un CRIS:

- 1) Diseño del Curriculum integral del investigador, que se refiere a la integración de diferentes perfiles requeridos por convocatorias institucionales, como son la Recategorización y Evaluación al Desempeño Docente, y nacionales, principalmente la del Sistema Nacional de Investigadores (SNI) y del Programa de Mejoramiento de Personal Académico (PRODEP).
- 2) Adopción de ORCID como identificador digital del investigador, perfil web público y conector con bases de datos externas como Web of Science o Scopus, para la recuperación automática de referencias de publicaciones y citas.
- 3) Interoperabilidad con el Repositorio Institucional NINIVE, con otras plataformas institucionales y otros sistemas CRIS.

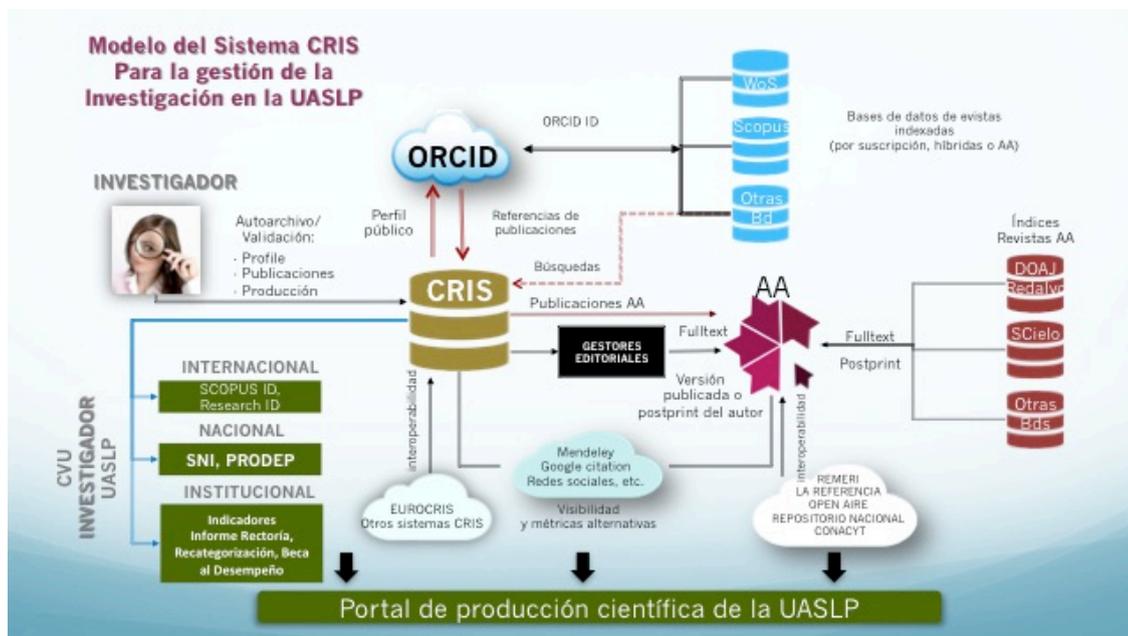


Fig. 2. Modelo del Sistema de Gestión de la Investigación de la UASLP.
 Fuente: Elaboración propia.

El desarrollo del modelo se planteo por etapas considerando los siguientes componentes estratégicos:

- Componente de gestión de contenidos:** Comprende la búsqueda y recuperación en Bases de Datos (Scopus, WoS, Redalyc y otras) de las publicaciones de los profesores de tiempo completo; la recopilación de los currículos o perfiles de profesores investigadores del SNI y/o con perfil PRODEP; el análisis de las políticas de revistas con gestores editoriales (SHERPA-ROMEEO, DULCINEA); y la creación de catálogos de autores, revistas, proyectos y fuentes de financiamiento.
- Componente tecnológico:** Comprende la implementación de la interfaz del investigador (*Back-end*) y la interfaz de salida (*Front-end*) del Sistema de Gestión de la Investigación; el desarrollo de un conector con el Repositorio Institucional NINIVE; la implementación de las API's de ORCID, SCOPUS y de otras bases de datos la API de ORCID y la generación de indicadores de producción y reportes de salida.
- Componente normativo:** Comprende el seguimiento y eventual aprobación por el H. Consejo Directivo Universitario del Mandato Institucional de Acceso Abierto; la correcta gestión de los derechos de autor a través del uso de licencias *Creative Commons* y el diseño de los procedimientos, flujos de trabajo y políticas de servicio.

Con la finalidad de obtener recursos financieros, alinearse a la Política Pública y

Decreto de Ley de Acceso Abierto²³ de México y formar parte del Repositorio Nacional de Recursos de Información Científica, Tecnológica y de Innovación²⁴, fue presentado en la convocatoria CONACYT 2015 para desarrollar los Repositorios Institucionales de Acceso Abierto, el proyecto “Desarrollo de un Modelo de Servicios para la consolidación del Repositorio Institucional NINIVE de la U.A.S.L.P.”, el cual obtuvo el 100% del financiamiento solicitado. La primera etapa del proyecto fue desarrollado durante el periodo agosto 2016 - agosto 2017, concluido en tiempo y forma con 23 productos o entregables. Uno de los resultados más importantes fue la implementación a nivel prototipo del Sistema de Gestión de la investigación llamado ORBIS, utilizando como componente *front-end* del sistema la plataforma *open source* VIVO v. 1.9.1. en el cual se capturaron los curriculums de 44 profesores de tiempo completo con sus publicaciones científicas del periodo 2012-2017 recuperadas de las bases de datos de SCOPUS y Web of Science.

Durante la segunda etapa agosto 2017- agosto 2018, se continuo trabajando en la parte técnica y de contenidos. Con el apoyo del equipo de documentalistas del Sistema de Bibliotecas se realizó la búsqueda y recuperación en las bases de datos de las publicaciones científicas del periodo 2012-2017 del total de los profesores investigadores de tiempo completo, mismas que fueron cargados a ORBIS junto con sus perfiles básicos. En cuanto a la implementación técnica, se diseño un método utilizando la herramienta KARMA para modelar los datos de las publicaciones a ontologías en RDF y cargarlas a la plataforma VIVO. Adicionalmente, se implementó la API institucional de ORCID, utilizando LDAP para autenticar la afiliación de los profesores y solicitar su autorización para leer y escribir en sus registros ORCID ID.

Para la tercera etapa, se presentó nuevamente en la Convocatoria CONACYT 2017 para desarrollar Repositorios de Ciencia Abierta, el proyecto “Implementación del Sistema de Gestión de Investigación (CRIS) para la mejora y poblamiento del Repositorio Institucional NINIVE”, el cual fue aprobado con el 100% de los recursos solicitados y fue llevado a cabo durante el periodo octubre 2018 – junio 2019. En esta etapa se llevaron a cabo las siguientes acciones estratégicas: 1) Implementación de la interfaz del investigador del Sistema ORBIS (*back-end*); 2) Desarrollo de un conector con NINIVE; 3) Configuración de ORCID con la nueva versión de NINIVE (DSpace v.6.2) y la certificación de la implementación; 4) Creación de tres catálogos de autores, revistas e índices, proyectos y financiamiento; 5) Capacitación a profesores sobre el registro ORCID, el auto archivo en NINIVE y la gestión de su hoja de vida en ORBIS.

En el siguiente apartado se describe la arquitectura del sistema con cada uno de los componentes de la solución tecnológica.

2. Descripción de la solución tecnológica implementada

Con base en el modelo descrito anteriormente y el trabajo que hasta el momento se ha

²³ DECRETO (DOF:20/05/2014) por el que se reforman y adicionan diversas disposiciones de la Ley de Ciencia y Tecnología, de la Ley General de Educación y de la ley Orgánica del Consejo Nacional de Ciencia y Tecnología. En: http://www.dof.gob.mx/nota_detalle.php?codigo=5345503&fecha=20/05/2014

²⁴ Repositorio Nacional. <https://www.repositorionacionalcti.mx/>

llevado a cabo, se diseñó la arquitectura del Sistema CRIS llamado ORBIS descrito en la Fig. 3., que comprende esencialmente cuatro grandes componentes:

1. Entradas: Constituyen las fuentes de información que alimentan al Sistema. Se consideran de dos tipos, las externas que comprenden las Bases de Datos de las cuales se exportan las publicaciones científicas tanto como comerciales como de Acceso Abierto; los registros ORCID ID de los investigadores, tanto para empujar la afiliación como para exportar publicaciones al Sistema ORBIS; y los catálogos nacionales e internacionales para la obtención y validación de identificadores digitales persistentes y vocabularios controlados para la descripción de ciertos metadatos. Las fuentes internas comprenden los catálogos institucionales de programas académicos, materias, dependencias y otros, así como también, el propio investigador quien valida y complementa la información cargada en el Sistema.
2. Salidas: Constituyen los medios para mostrar la información de las hojas de vida de los investigadores y su producción científica. Son tres interfaces: el portal público o componente *front-end* de ORBIS basado en el software VIVO v. 1.9.2., el módulo de reportes de la interfaz del investigador para generar su curriculum con base en determinados perfiles, y la generación de indicadores de producción científica para diferentes niveles y perfiles de usuario.
3. Sistema de gestión del curriculum: Es propiamente el sistema CRIS, también referido como el componente *back-end* de ORBIS, que comprende diferentes módulos para la actualización del curriculum del investigador, la generación de reportes por perfiles de usuario y las herramientas para la importación y exportación de información.
4. Interoperabilidad: Basicamente se refiere a la conexión con el Repositorio Institucional NINIVE basado en el software Dspace v.6.2, a través de la configuración del protocolo SWORD, de tal forma que el investigador no tenga que registrar dos veces una misma publicación, sino que al registrarla en ORBIS, se transfiera el documento y sus metadatos de manera automática a NINIVE mediante este protocolo.

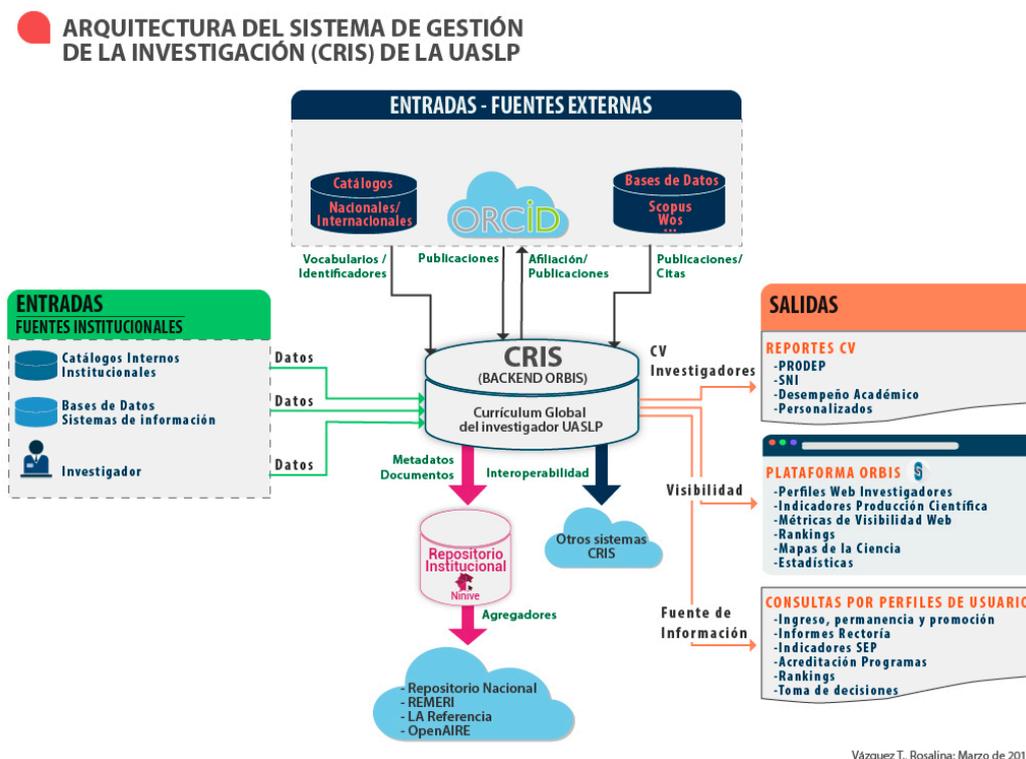


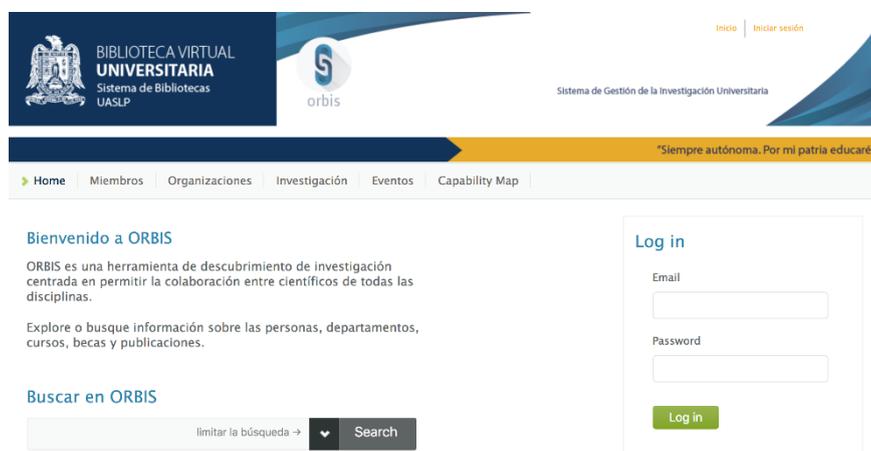
Fig. 3. Arquitectura del Sistema CRIS de la UASLP. Fuente: Elaboración propia

En los siguientes apartados se describe la implementación de las tecnologías abiertas empleadas para el desarrollo de los principales componentes del modelo propuesto.

2.1 Plataforma de software VIVO

VIVO es un software open source que forma parte del grupo DuraSpace, basado en ontologías en RDF para la representación de las publicaciones académicas. Algunas de las funcionalidades de VIVO son el registro, edición, búsqueda, navegación y visualización de información sobre investigadores, organizaciones, actividades y otras entidades. Permite enlazar mediante el modelo de datos ligados (Linked data) los perfiles de los investigadores con su producción científica y las organizaciones de afiliación, financiamiento u otras [9].

Este software es utilizado para la implementación del portal público o componente *front-end* del Sistema ORBIS, disponible en el sitio: <http://orbis.uaslp.mx/vivo/>.



2.2 ORCID

ORCID es una organización sin fines de lucro que ofrece un sistema para crear y mantener un registro único de investigadores y vincular sus actividades y productos de investigación. Constituye también un perfil web público y un medio de conexión con bases de datos de publicaciones. Actualmente es el sistema de identificación de autores más reconocido y utilizado en la publicación científica [10].

La implementación de la API institucional de ORCID comprende la realización de las siguientes tareas:

- a) Configuración de la aplicación:
<https://bv.uaslp.mx/afiliacion/index.html>



- b) El investigador ingresa sus credenciales universitarias y por medio de LDAP se realiza una consulta a la base de datos de empleados para verificar que el usuario pertenezca a la Institución.
- c) La aplicación devuelve los datos del investigador y le solicita seleccione su nombramiento y año de ingreso. Enseguida selecciona el botón *Create or Connect your ORCID ID*
- d) Si ya tiene una cuenta de ORCID ingresa con sus credenciales, de lo contrario, selecciona la opción de *Registrarse* para crearla
- e) El investigador autoriza a la UASLP leer/escribir en su ORCID ID
- f) El ORCID ID queda vinculado a la UASLP y se empuja la afiliación al registro

2.3 Sistema de gestión del curriculum

El sistema está desarrollado con los lenguajes HTML, CSS, JavaScript y PERL. El sistema de base de datos relacional está diseñado en PostgreSQL.

Los módulos del sistema son los siguientes:

- Autenticación: Interfaz para realizar el inicio de sesión al sistema
- Investigador: interfaces donde el investigador puede consultar y modificar su perfil, así como generar reportes sobre su perfil/CV.
- Administrador: Interfaces para realizar tareas de importación y exportación de información.
- Perfiles de usuario: Interfaces para diferentes funcionarios y autoridades que requieran consultar información y/o generar reportes e indicadores de producción científica.

Algunas de estas interfaces se muestran en las siguientes capturas de pantalla:

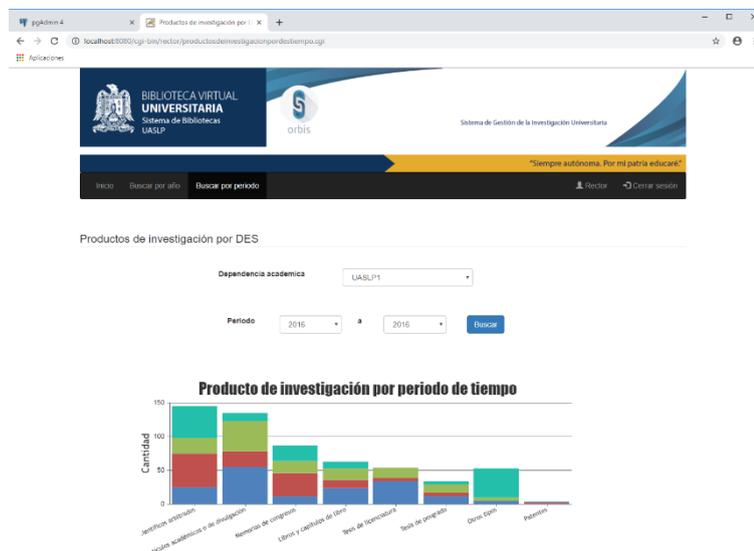
Interfaz de datos generales del investigador

The screenshot shows a web browser window displaying the 'Datos Generales' form. The browser address bar shows 'localhost:8080/cgi-bin/investigador/datosGenerales.cgi'. The page header includes the UASLP logo and the text 'Sistema de Gestión de la Investigación Universitaria'. The form is titled 'Información general' and contains the following fields:

CURP:	RFC:	Fecha de nacimiento:
Nombre:	Apellido Paterno:	Apellido Materno:
Sexo:	Estado conyugal:	País de nacimiento:
Entidad federativa:	No de CVU:	Nacionalidad:

Below these fields, there is a section for 'IES de adscripción' with a dropdown menu showing 'Universidad Autónoma de San Luis Potosí'. At the bottom, there is a section for 'Áreas y disciplinas'.

Interfaz de generación de reporte del perfil del Rector



2.4 API's de Bases de datos

Para la recuperación de las publicaciones científicas de los investigadores de la UASLP se han considerado hasta el momento las siguientes bases de datos:

- SCOPUS: Las APIs de Scopus se basan en servicios REST y son de uso gratuito; sin embargo, es necesario contar con una suscripción a Elsevier. Para comenzar a usar la API, el primer paso es registrarse para obtener una API Key (<https://dev.elsevier.com/index.jsp>). Enseguida, se debe seleccionar la API que se desea configurar y seguir el procedimiento indicado en el sitio.
- CROSSREF: La API está disponible de manera pública desde el sitio (<https://github.com/CrossRef/rest-api-doc>) por lo que no requiere de ningún registro previo. Para configurarla, las llamadas se hacen a través del protocolo HTTPS y se obtiene una lista con la información de todas las publicaciones de un autor. También permite hacer peticiones usando el ORCID del autor, en este caso utiliza el DOI del artículo como identificador para cada publicación.
- PUBMED: Al igual que en Scopus se requiere hacer un registro para obtener el permiso de uso de la API. En el sitio web (<https://www.ncbi.nlm.nih.gov/home/develop/api/>) se registra la aplicación para obtener una API Key y además se hace un registro enviando un correo a utilities@ncbi.nlm.nih.gov con el nombre del desarrollador encargado de la conexión, la organización, el nombre de la aplicación, y el correo de contacto. Una vez completado el registro se tiene un límite de frecuencia de hasta 10 peticiones por segundo.

2.5 Protocolo SWORD

El protocolo SWORD (Simple Web-service Offering Repository Deposit) permite depositar documentos y sus metadatos en un Repositorio Institucional desde una fuente externa. Para DSpace v.6.2 están disponibles varias versiones para su implementación. Una de las ventajas de usar SWORD es que cuenta con librerías en una gran variedad de lenguajes lo que facilita su implementación. En nuestro caso, se utiliza la librería de SWORD en PERL disponible en el sitio <https://github.com/zostay/Sword-perl>, para depositar una publicación registrada en el Sistema ORBIS, en el Repositorio Institucional NINIVE.

2.6 Herramientas D2R y SPARQL

La plataforma D2RQ es un sistema para acceder a bases de datos relacionales como si fueran grafos RDF, ofrece acceso basado en RDF al contenido de las bases de datos sin tener que replicar la información en una RDF store. La plataforma D2R incluye la herramienta *d2r-query*, la cual permite realizar consultas SPARQL sobre la base de datos relacional del Sistema CRIS y devuelven los datos con el formato correcto para ser insertados en la plataforma ORBIS (VIVO) mediante la API "*SPARQL Update APP*".

3. Resultados obtenidos e impacto

A tres años de haber iniciado el proyecto, se han obtenido los siguientes logros y beneficios:

- En las dos ocasiones que fue presentado el proyecto en las convocatorias públicas de CONACYT, fue aprobado con el 100% financiamiento solicitado, lo que permitió, en la tercera etapa 2018-2019, la contratación de cuatro becarios de licenciatura de las carreras de Gestión de la Información y de Ingeniería en Computación, así como la contratación por honorarios de un programador que se dedicó exclusivamente al desarrollo del componente *back-end* del Sistema ORBIS.
- Hasta el momento (mayo de 2019), están capturados en la plataforma ORBIS (<http://orbis.uaslp.mx/vivo/>), los perfiles básicos del total de profesores investigadores de tiempo completo de la UASLP (946) de los cuales, el 10% ya tiene capturado su curriculum completo. En cuanto a la producción científica se tiene el acceso a 3,051 artículos publicados correspondientes a 1,494 revistas arbitradas, 83 libros, 57 capítulos de libro, 126 publicaciones en Conferencias, 8 patentes, 4 proceedings y 28 premios, entre otros. También se encuentran registrados los cuerpos académicos (grupos de investigación) con sus líneas de generación y aplicación del conocimiento, así como también, los organismos financiadores vinculados a las publicaciones.
- En cuanto a ORCID, se han autenticado a través de la API institucional,

aproximadamente un 25% de los investigadores. Durante el presente año, se continuará con la organización de talleres y webinars para sensibilizar a los investigadores sobre los beneficios de crear y/o vincular su registro ORCID ID y proporcionarles los materiales de capacitación necesarios, mismos que son publicados en el micro sitio de ORCID desde el sitio de la Biblioteca Virtual Universitaria (BVU): <https://bvu.uaslp.mx> y en plataforma educativa TZALOA de la BVU (<https://tzaloe.uaslp.mx>)

- Respecto al desarrollo del Sistema de gestión del curriculum, que constituye el componente back-end del ORBIS, se llevará a cabo hacia finales de año, una segunda prueba de usabilidad con un grupo piloto de investigadores, con la finalidad de complementar y mejorar la funcionalidad del sistema. Adicionalmente, se tiene contemplado implementar algunas otras API's de bases de datos y herramientas de métricas alternativas.
- El proyecto cuenta con el respaldo de la Rectoría y del acuerdo de la gran mayoría de los directores y funcionarios a cargo de las dependencias académicas, lo que constituye una importante ventaja para el desarrollo del proyecto. Sin embargo, el proyecto aún no está consolidado lo que vuelve vulnerable su continuidad ante el cambio de Rector en abril de 2020.

Adicionalmente, como resultado de los aprendizajes se plantean las siguientes conclusiones generales:

- Hay una tendencia creciente a desarrollar sistemas CRIS e integrarlos a los repositorios para construir ecosistemas de investigación en el marco de la ciencia abierta.
- Actualmente en EuroCRIS están registrados 210 sistemas CRIS, soportados mayormente sobre software abierto o desarrollo propio.
- Cada vez más se integra ORCID a los repositorios y a los CRIS, para proveer identificadores digitales persistentes a los investigadores y crear sus perfiles web.
- No hay una metodología única para implementar un CRIS. La solución tecnológica dependerá de las necesidades, características y recursos disponibles de la Institución, pero sobre todo, de lo que se quiera lograr en el mediano y largo plazo.

Referencias

1. European Commission: Digital Science in Horizon 2020. European Commission (2013)
2. Bernal Martínez, I, Vallés, PT, García-Puente, M, Fundadoras De Socialbiblio, S: Tendencias y cuestiones en sistemas de información científica. 70,7 (2014)
3. Bryant, R, Clements, A, Groenewegen, D, Huggard, Mercer, H, Missingham, R, Oxnam, M, Rauh, A, Wright, J: Research Information Management: Defining RIM and the Library's Role (2017)
4. Red de Bibliotecas Universitarias Españolas REBIUN: Sistemas CRIS y Repositorios Institucionales en las Universidades Españolas (2013).
5. DSpace-CRIS Home - Dspace-CRIS – DuraSpace Wiki,

- <https://wiki.duraspace.org/display/DSPACECRIS/Dspace-CRIS+Home>
6. OpenAIRE: The OpenAIRE Guidelines for CRIS Managers version 1.1. (2018)
 7. Lynch, C: Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age ARL. A Bimonthly Report, no. 226 (2003)
 8. Confederation of Open Access Repositories COAR: Incentivos, integración y mediación: Prácticas sostenibles para poblar repositorios (2013)
 9. VIVO – DuraSpace Wiki, <https://wiki.duraspace.org/display/VIVO>
 10. ORCID, <https://orcid.org/>

Capítulo 5

Smartcampus y Tendencias – Blockchain

SIU-Guaraní y la Universidad Digital: Registro de información académica con Firma Digital y Blockchain

Hernán Turin, Guillermo Diorio, Guillermo Trutner, Sebastián Natarelli

Sistema de Información Universitaria (CIN-SIU), Ecuador 871 Piso 3 CABA - C1214ACM - Argentina
hturin@siu.edu.ar, gdiorio@siu.edu.ar, gtrutner@siu.edu.ar, snatarelli@siu.edu.ar

Resumen: La transformación digital ha dejado de ser una opción para las organizaciones, incorporar nuevas tecnologías e impulsarlas hacia su interior, se ha convertido en una necesidad vital. Las Instituciones universitarias no son ajenas a esta realidad, por el contrario, tienen el desafío de incorporar y liderar estos cambios organizacionales. Deben implementar una administración electrónica integral, con procesos completamente digitales, que les permitan optimizar el uso de sus recursos y dar una rápida respuesta a los requerimientos de alumnos, docentes y sociedad en general. La suite de soluciones desarrolladas por el SIU, basadas en software libre, acompañó a las instituciones argentinas con herramientas tecnológicas robustas e innovadoras, que propiciaron la adopción de estos nuevos procesos. Este conjunto de soluciones, permiten hoy contar con una administración completamente electrónica que incorpora, entre otras cosas, firma digital, un repositorio institucional de documentación digital, y el uso de blockchain para brindar trazabilidad y seguridad a diferentes procesos organizacionales.

La gestión académica es clave en los procesos de modernización. SIU-Guaraní provee un mecanismo de gestión de actas y registros académicos que desde sus inicios ha presentado niveles de seguridad altos. Ésta última versión permite asegurar la integridad y seguridad de la información, implementando un circuito que permite firmar digitalmente el documento electrónico y almacenarlo en un repositorio único. Además, el conjunto de datos que componen el acta se almacena en blockchain, lo que posibilita asegurar la integridad de los datos almacenados en el sistema de gestión académica.

Palabras claves: Sistema de Información. Gestión. Servicios. Firma Digital. Blockchain.

Ejes Temáticos: Logrando universidades digitales seguras. Smartcampus y tendencias tecnológicas emergentes en la Universidad.

Introducción

El Sistema de Información Universitaria de la República Argentina (SIU) fue creado por la Secretaría de Políticas Universitarias del Ministerio de Educación de la Nación en 1996. Desde 2013 forma parte del Consejo Interuniversitario Nacional (CIN) cuyas funciones son, esencialmente, la coordinación y promoción de políticas universitarias y actividades de interés para el sistema público de Educación Superior [1] y es, además, órgano de consulta obligada para la toma de decisiones de trascendencia para cualquier ámbito del sistema universitario nacional.

El SIU desarrolla sistemas para la toma de decisiones, el análisis institucional y la gestión en el ámbito de las Universidades Nacionales con el objetivo de dotar al sistema de Educación Superior de elementos que permitan mejorar la confiabilidad, completitud, disponibilidad e integridad de la información. Los sistemas SIU son diseñados, desarrollados e implementados con criterios homogéneos y consensuados por quienes serán sus usuarios, y ofrecen soluciones para las distintas áreas (académica, económico-financiera-presupuestaria, recursos humanos, bibliotecas, compras y contrataciones, patrimonio, facturación y cobranzas etc.) y su correspondiente módulo de análisis para uso gerencial.

La premisa imperante del SIU es que la tecnología debe estar al servicio de las instituciones, lo cual implica acompañar los objetivos institucionales considerando a las personas como el eje central.

La estructura interna de la organización está conformada por una Coordinación General y una segunda línea de coordinadores que abordan temáticas técnicas y funcionales, la gestión académica, de recursos humanos o temas gerenciales, solo por citar algunos ejemplos. Existen además áreas transversales que se encargan de temáticas comunes como ser Bases de Datos, Análisis Funcional, Testing, etc. Una vez al mes se realiza una reunión de la que participan coordinadores de proyectos y áreas, en la que se comparten objetivos y estrategias comunes, metas, problemáticas y casos de éxito.

El SIU-Guaraní

Una visión holística del proyecto permite vislumbrar que el sistema de gestión académica SIU-Guaraní integra componentes sociales, tecnológicos, políticos, culturales y económicos que interactúan entre sí. Su objetivo es desarrollar un único sistema informático para todas las Instituciones Universitarias de gestión pública de Argentina, con una arquitectura técnica que les permita extenderlo o personalizarlo según sus necesidades particulares manteniendo la compatibilidad. Estas características permiten al sistema reflejar la realidad de cada una de las instituciones.

Por otro lado, con el objeto de mejorar los procedimientos para hacerlos más eficientes, paulatinamente se incorporan mejores prácticas en el sistema que introducen mejoras en la toma de decisiones, el análisis institucional y la transparencia de la información. Estas son sólo algunas de las características que convierten al proyecto en lo que podemos denominar autosustentable.

Desde la visión de producto o sistema informático, SIU-Guaraní es un software que

administra la gestión académica desde que los alumnos ingresan como aspirantes hasta que obtienen el diploma. Permite llevar adelante la gestión de alumnos en forma segura e íntegra, suministrando información consistente y oportuna para la toma de decisiones. Además provee mecanismos que permiten su adaptación a los procesos de gestión académica, oferta educativa y estructura organizativa de cada Institución.

El contexto y los avances

La incorporación de SIU-Guaraní en las universidades nacionales representó el comienzo de un cambio en los paradigmas y los procesos internos que las instituciones llevaban adelante en cuanto al manejo de información. Uno de los principales procesos afectados fue el de registro de la actuación académica de un alumno, en el que se han incorporado modificaciones que permitieron optimizarlo en forma sustancial (Fig. 2).

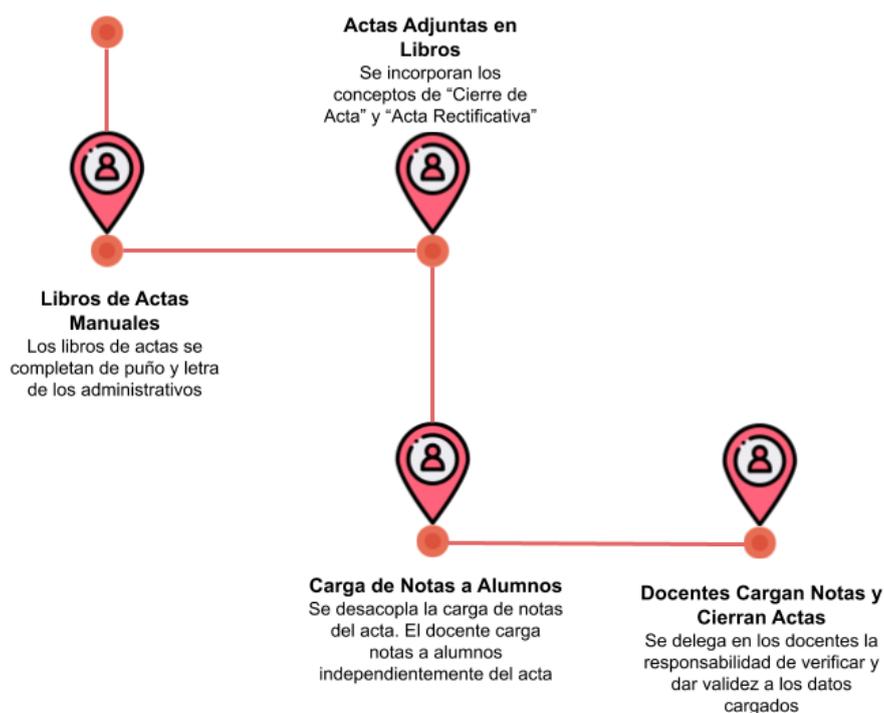


Figura 2 - Evolución del circuito de actas

Las instituciones universitarias argentinas han registrado, históricamente, la actividad académica de los alumnos en libros de actas. Estos libros son un compendio de actas generadas y rubricadas por los docentes que contienen uno a uno los registros de la actividad académica de cada alumno, ya sea por la aprobación de un examen, la

promoción de una actividad o la aprobación de la cursada de la misma.

El arribo de SIU-Guaraní, hacia 1999, y los mecanismos de seguridad con los que el sistema contaba en ese momento, traccionaron los primeros cambios en el proceso. Las instituciones comenzaron a adquirir confianza en la tecnología, y muchas de ellas decidieron dejar de transcribir los registros en forma manual a los libros de actas preimpresos, y comenzaron a generarlos a partir de la encuadernación de las actas impresas por el sistema. En algunos casos se trataba de un formulario impreso con los datos mínimos de los alumnos, donde los resultados obtenidos por cada persona eran completados de puño y letra del docente. En otros se adjuntaba una impresión del acta con los datos registrados en el sistema que contenía, al pie de cada página, la firma docentes involucrados.

Uno de los mecanismos que brindó mayor seguridad al proceso fue la incorporación de una instancia de validación y control denominada "Cierre de Actas" cuyo objetivo era que los empleados administrativos de las áreas de alumnado pudieran verificar, en conjunto con el docente, la completitud y correctitud de los datos registrados en el sistema para, luego de esto, otorgarle validez legal a la actuación (Fig. 3).

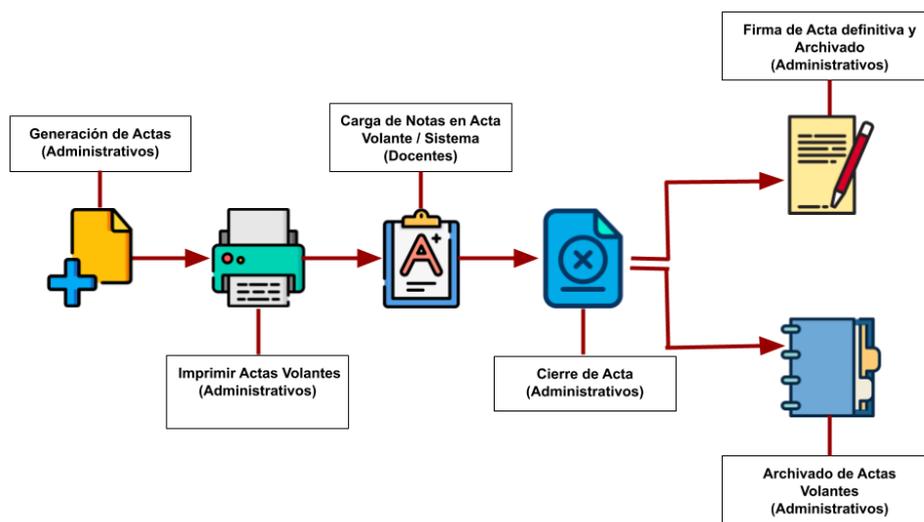


Figura 3 - Circuito de actas con acta volante

El avance de la tecnología y la apropiación de la solución por parte de los diferentes actores dentro de las universidades, permitió avanzar en la implementación de mejoras en este circuito. Hasta ese momento los docentes sólo podían registrar notas en actas previamente generadas por el personal administrativo. El sistema les brindó la posibilidad de registrar las notas de cada uno de sus alumnos, independientemente de si el acta estaba previamente generada. Esta información provista por el docente, se utilizaría luego como insumo para el proceso de generación de actas.

A la luz los buenos resultados obtenidos en las mejoras implementadas y, luego de superar algunos aspectos relacionados a la idiosincrasia de las instituciones, el paso siguiente fue delegar el proceso de "Cierre" (verificación, validación y consiguiente validez de la actuación consignada en el acta) a los docentes, lo que representó uno de

los grandes saltos cualitativos en el circuito y permitió la reducción de los tiempos necesarios para dar como válida una actuación académica de una persona en particular (Fig. 4). Este nuevo esquema, sin embargo, sigue teniendo como principal protagonista al acta en papel firmada de puño y letra del docente.

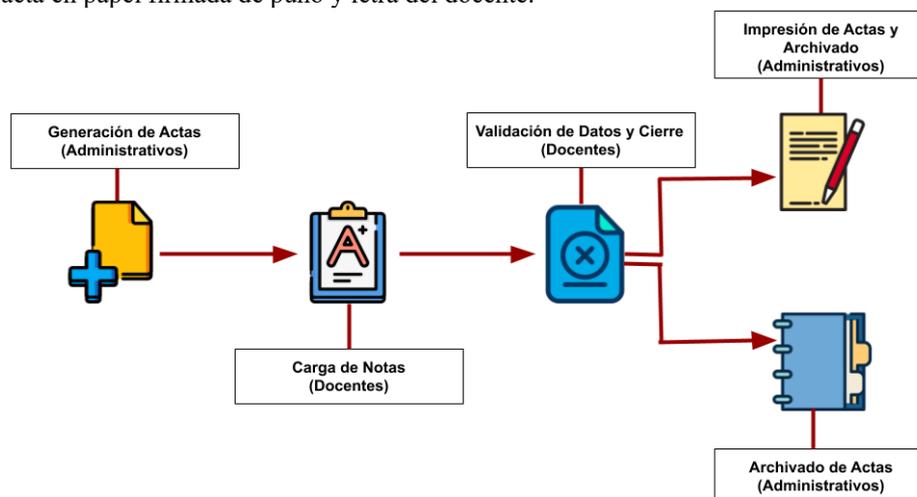


Figura 4 - Circuito de actas con cierre del docente

La Firma Digital

En Argentina, el advenimiento de la firma digital se da en el año 2001 con la sanción ley 25.506 [2] cuya reglamentación se llevó a cabo a través del decreto 2628 de 2002 [3] que establece las características técnicas que debe cumplir una firma para tener la categoría de digital. Sin embargo, no fue hasta 2015, cuando el tema se incorporó al nuevo Código Civil, que obtuvo la misma validez legal que la que tiene la firma holográfica.

En la actualidad existen dos formas aceptadas y legalmente válidas para realizar la firma digital: Firma Digital por Hardware y Firma Digital Remota [4]. La primera requiere de un token en poder del firmante, mientras que la segunda no necesita de un dispositivo físico. La firma digital remota se puso en marcha en 2018 y simplifica el despliegue de la firma dentro de las instituciones, ya que se puede hacer desde cualquier dispositivo, no tiene asociado el costo de compra de un token ni requiere la instalación de drivers en las PC o dispositivos de trabajo desde donde se realice la firma.

En 2013 se desarrolló un firmador digital que posteriormente integramos a las aplicaciones de SIU. Como primera aplicación se incorporó la firma en los reportes de información estadística de alumnos que las Universidades envían año a año al Ministerio de Educación de la Nación. En 2014 incorporamos la firma digital en los recibos de haberes generados a partir de la liquidación de sueldos realizada por el sistema de gestión de recursos humanos SIU-Mapuche.

A pesar de estos intentos su implementación en el sistema universitario ha sido bajo, entre otras cosas por la complejidad de transformar a las Universidades en

Autoridades de Registro necesarias para el otorgamiento de la FD, por el alto costo de los dispositivos Token y por la complejidad de mantener un entorno controlado en los equipos informáticos de quienes realizan la firma.

El avance de la administración electrónica y la digitalización de los procesos que permiten el intercambio de información entre los diferentes organismos en formatos completamente electrónicos, hacen que se vuelva imprescindible implementar la firma digital para brindar seguridad y validez a la información que circula en el mundo digital.

Blockchain

El avance de blockchain a nivel internacional ha demostrado su enorme capacidad para garantizar seguridad y transparencia en diferentes tipos de procesos. Estas cadenas de bloques surgieron como una tecnología distribuida, donde hoy se pueden registrar transacciones, las cuales se encuentran distribuidas a lo largo de la red asegurando la transparencia, la autoría y la inalterabilidad de las mismas sin intermediarios.

En Argentina se está trabajando desde el año 2018 en el desarrollo de Blockchain Federal Argentina, que es la primera plataforma multiservicios de alcance federal basada en blockchain [5]. En esta iniciativa colaboran NIC Argentina, la Cámara Argentina de Internet - CABASE - y la Asociación de Redes de Interconexión Universitaria (ARIU).

A través de esta iniciativa conjunta, en la que las partes representan al sector público, el sector privado y la academia, se conformó la infraestructura sobre la que corre la primera plataforma nacional de uso público basada en Blockchain, una innovadora tecnología de validación de transacciones.

Esta plataforma multiservicios es gratuita, transparente, segura y auditable, y fue concebida para servir como base para desarrollos de toda la comunidad. Además se planea como un sistema abierto, que se pueda fortalecer con los aportes de los diferentes sectores que componen el ecosistema de internet.

Como parte de los beneficios, BFA asegura la transparencia dado que la información que se sube a BFA es pública, visible y completamente auditable. Además de la confiabilidad dado que utiliza un modelo robusto y descentralizado, se asegura un entorno resiliente y garantiza la seguridad de la información. En este esquema tampoco hay intermediarios, esto se logra gracias a que la información guardada en la cadena es segura por naturaleza.

Otro de los beneficios, es la ausencia de uso de criptomonedas. BFA no necesita una moneda virtual para aprovechar las ventajas de blockchain. Y finalmente, BFA opta por un modelo que habilita un bajo costo tanto en procesamiento como consumo de energía y ofrece transacciones sin costo.

En el SIU, desde mediados de 2018, se está trabajando en el uso de esta tecnología y plataforma y su aplicación para los diferentes desarrollos. El primero en integrarse a BFA fue el Registro Público de Graduados Universitarios [7]. Este registro contiene la información de todos los graduados universitarios de Argentina. Con esta implementación, se puede validar online la veracidad de la información de este registro (Fig 5).

REGISTRO PÚBLICO DE GRADUADOS UNIVERSITARIOS

SISTEMA INFORMÁTICO DE CERTIFICACIONES

Nombre y apellido: dos Reis, Wanderiel José

Documento	Apellido	Nombre	Nacionalidad	Institución	Título	Egreso	Registro BFA
Cédula de Identidad	dos Reis	Wanderiel José		Universidad	Doctor en Derecho	02/07/2014	Detalle

CONVALIDACIONES DE TÍTULOS UNIVERSITARIOS EXTRANJEROS

No se encontraron resultados en el sistema de Convalidaciones de títulos universitarios extranjeros.

Desarrollado por SIU-CIN
2002-2019

Figura 5 - Registro Público de Graduados Universitarios integrado con BFA

Plataforma Integradora de Servicios: SIU-Araí

Hace varios años el SIU comenzó a trabajar en el desarrollo de SIU-Araí, una plataforma que permite integrar las diferentes soluciones propias y externas haciendo simple y transparente su integración. El resultado es una herramienta genera un ecosistema de aplicaciones que interactúan entre sí consumiendo y ofreciendo servicios. Gracias a su puesta en marcha, que propicia la utilización de estándares, posibilita la integración con aplicaciones y servicios de terceros de forma simple y ordenada. La figura 6 expone el detalle de la arquitectura general de la plataforma, así como la vinculación entre componentes internos, módulos del SIU y aplicaciones o sistemas externos.

La plataforma cuenta, hasta el momento, con cuatro componentes: Araí-Registry (Registro de aplicaciones y servicios), Araí-Usuarios (Proveedor y Administrador de Identidad para toda la plataforma) que además provee el servicio de ingreso único (Single Sign On), Araí-Cli (librería de cliente de la plataforma) que provee de servicios de bajo nivel facilitando la integración de los componentes y Araí-Documentos, que tiene dos funcionalidades esenciales: Repositorio Digital y Bandeja de Autorización.

El Repositorio Digital, herramienta destinada a alojar documentos digitales, está diseñado para integrarse con todos los sistemas SIU de manera nativa y permite la integración de cualquier sistema externo gracias a su tecnología REST. Algunas de sus principales funcionalidades son:

- Dejar los documentos disponibles en la bandeja de autorización/firma.
- Informar al sistema de origen cuando un documento es autorizado/firmado.
- Agrega de manera automática, luego de ser firmados/autorizados, los documentos que forman parte de un trámite.

- Identifica los documentos por "tipo" y les asigna permisos para dejarlos disponibles para los trámites y áreas que puedan necesitarlos.

Por su parte, la Bandeja de Autorización gestiona cada uno de los documentos que debe ser autorizado por la persona correspondiente. Se entiende por autorización a la Firma Digital, Firma Electrónica o el Visado de documentos. Los permisos de autorización son a nivel persona, no de área.

Otro de los aspectos interesantes de SIU-Araí es el Portal del empleado SIU-Huarpe, que oficia de único acceso para los usuarios y concentra la totalidad de los servicios ofrecidos por la institución. Un detalle importante es que el portal no cuenta con funcionalidades propias sino que la mayoría de sus servicios son provistos por el resto de las soluciones SIU

Integración de SIU-Guaraní con SIU-Araí

Con la puesta en marcha de SIU-Araí, las diferentes soluciones del ecosistema SIU debieron transformarse para poder integrarse a esta plataforma, y SIU-Guaraní no fue la excepción. El primer paso en este proceso fue la integración con el componente Araí-Usuarios. Si bien desde el año 2014 SIU-Guaraní tenía implementada la posibilidad de utilizar Proveedor de Identidad (IdP), con el surgimiento de la nueva plataforma integradora, el primer objetivo fue brindar la posibilidad de autenticar a las personas utilizando Araí-Usuarios. Este fue un paso clave ya que es la puerta de entrada para luego poder acceder al resto de los beneficios provistos por SIU-Araí.

En este proceso, el segundo paso fue la implementación del componente Araí-Documentos. Este es uno de los componentes más nuevos de la plataforma. El objetivo del mismo es brindar un repositorio de documentación digital único para toda la institución y además proveer de una bandeja de documentos para la firma digital de los mismos. Si bien SIU ya ofrecía un Repositorio Digital Institucional para universidades (implementado en Nuxeo), la decisión fue que este repositorio pase a formar parte de SIU-Araí como parte de un nuevo componente denominado Araí-Documentos.

Validez Legal e Inalterabilidad: Firma Digital + Blockchain

Alineados con una de las premisas principales que rigen al SIU desde sus inicios, dotar a las universidades de herramientas que les faciliten la implementación de una administración electrónica integral, se plantea un nuevo desafío para el módulo de gestión de alumnos: incorporar una nueva tecnología que permita prescindir de documentos en papel para el registro de la actuación académica de un alumno. Avanzar en este sentido implicaba abordar dos aristas muy importantes: reemplazar las actas en papel con firma holográfica manteniendo su validez legal y por otro lado asegurar que los datos registrados en el sistema de gestión no se han adulterado a lo largo del tiempo.

La solución planteada para esta problemática surge desde dos tecnologías diferentes que se integran en una misma solución: la validez legal de las actas se da con la utilización de firma digital de un documento PDF, mientras que el control de la

inalterabilidad de los datos almacenados en el sistema estará dado por la utilización de Blockchain que, en este caso, se implementa con Blockchain Federal Argentina (BFA) (figura 7).

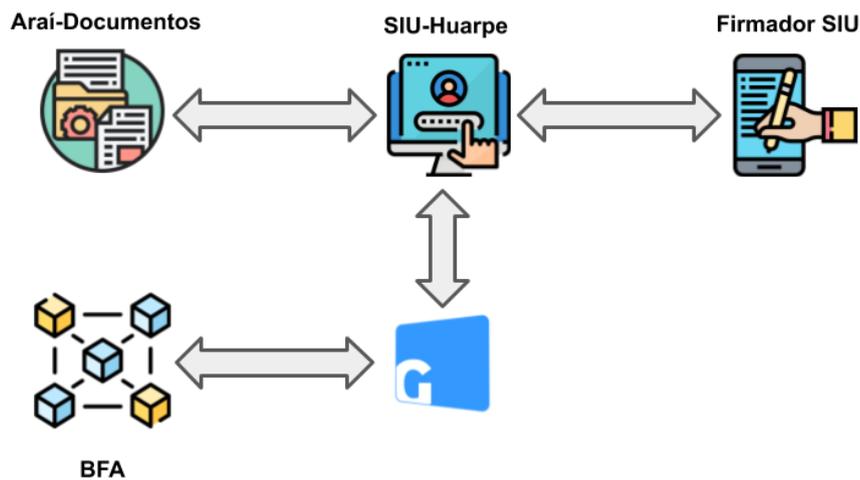


Fig. 7 - Esquema de solución diseñada

En el armado de esta solución, una de las primeras decisiones a tomar fue definir qué información del acta se enviará, convertida en un hash, a BFA para su resguardo. Esta definición implicó analizar cuál era el objetivo de almacenar las actas en el blockchain. Dado que la validez legal del acta está asegurada a través de la implementación de la firma digital, no es necesario enviar a blockchain la totalidad de los campos del acta con sus identificadores y sus descripciones, sino que para este objetivo será suficiente el envío de los identificadores más los campos específicos del detalle del acta (figura 8). Esta es información suficiente y necesaria para validar que la información contenida en el registro no ha sido alterada.

Para realizar el envío de la información a BFA, se definió un conjunto de campos del acta que son los que se utilizarán para construir el hash. Los campos a enviar son: id del acta, número del acta, tipo, origen, código de la actividad, nombre de la actividad, acta de referencia, y lista de alumnos con tipo de documento, número de documento, apellidos, nombres, instancia, fecha, nota y resultado. Estos campos se incluyen en un JSON (formato de texto sencillo utilizado para el intercambio de datos) y se le aplica un algoritmo, en nuestro caso utilizamos sha256, para construir el hash a enviar a Blockchain.

A continuación se muestra un ejemplo del JSON:

```
{
  "id": "202578",
  "numero_acta": "66321",
  "tipo": "Normal",
  "origen": "Examen"
  "actividad_codigo": "INGV"
  "actividad_nombre": "Taller de Inglés Nivel III"
  "acta_referencia": " "
  "alumnos": [ { "tipo_documento": "DNI",
                  "nro_documento": "27294565"
                  "apellidos": "Perez"
                  "nombres": "Roman Javier"
                  "instancia": "Examen Libre"
                  "fecha": "2010-08-06"
                  "nota": "7"
                  "resultado": "Aprobado" }
                ]
}
```

Adicionalmente se modificó el modelo de datos en SIU-Guaraní en el que se agregó una nueva tabla denominada `sga_actas_detalle_blockchain` (Tabla 1) en la que se almacenan el JSON generado, el hash enviado a blockchain, el recibo enviado por blockchain luego del registro.

Tabla 1 - Estructura de la tabla `sga_actas_detalle_blockchain`

Campo	Tipo de dato	Permite nulos
<code>json_original</code>	jsonb	NO
<code>hash_original</code>	text	NO
<code>block_number</code>	text	SI

La firma digital se implementa utilizando Araí-Documentos, componente que provee el repositorio digital en el que se almacenan los documentos digitales y, además, provee una bandeja en la que el docente puede ver cuáles son los documentos que tiene para la firma. esta visualización se realiza desde el portal SIU-Huarpe. El firmador propiamente dicho, no es parte del SIU-Araí, pero interactúa con él. Llegado este punto contábamos con dos alternativas para que el docente realice la firma, desde el portal SIU-Huarpe o desde el portal de autogestión de SIU-Guaraní, desde el que realizan las gestiones relacionadas a comisiones y exámenes. Para la primera etapa de trabajo se tomó la decisión de realizarla desde SIU-Huarpe, en una segunda etapa estará disponible también desde el portal de autogestión.

El docente puede firmar los documentos que tiene disponibles y, una vez firmado, se envía a BFA un hash generado a partir de los datos contenidos en el JSON, y el sistema SIU-Guaraní almacenará el recibo correspondiente al registro en blockchain (Fig 9). Para asegurar que la información de un acta no se envíará y registrará varias veces a BFA con adulteraciones, se desarrolló un “Smart Contract” (contrato inteligente) que tiene inteligencia necesaria para evitar que se registre más de una vez la misma acta. Para lograr esto, al enviar la información a BFA, no sólo se envía el

hash del acta, sino que además se envía otro hash con información que permite identificar en forma unívoca el acta. Este segundo hash se obtiene utilizando un JSON que contiene el código de SIU-Araucano de la Institución (código único que identifica a la Institución), el código de la Responsable Académica (Facultad) y el identificador del acta, al que se le aplica el algoritmo SHA256. Con esta información, el smart contract podrá verificar, cada vez que se quiera registrar un acta en BFA, si la misma ya ha sido registrada y, en caso de ya estarlo, evitar un segundo registro.

sga_actas

Columns

Key	Column name	Domain	Data type	Not null	Unique
PK	id_acta		Serial	YES	NO
	nro_acta		Varchar (30)	YES	NO
FK	origen		Char (1)	YES	NO
	tipo_acta		Char (1)	YES	NO

sga_actas_detalle

Columns

Key	Column name	Domain	Data type	Not null
	fecha_generacion			
	fecha_cierre			
	nua			
FK	documento			
	version			
	version_impresa			
	nro_ultima_copia			
	renglones_folio			
	acta_referencia			
	fecha_anulacion			
	observaciones			
FK	estado			
	cerrada_por_docente			
	id_acta		Integer	YES
PFK	alumno		Integer	YES
FK	instancia		Smallint	YES
FK	plan_version		Integer	YES
	fecha		Date	NO
	fecha_vigencia		Date	NO
	folio		Smallint	YES
	renglon		Smallint	YES
	pct_asistencia		Numeric (5,2)	NO
FK	escala_nota		Integer	NO
FK	nota	codigo 10	Varchar (10)	NO
FK	resultado		Char (1)	NO
FK	cond_regularidad		Integer	NO
	rectificado		Char (1)	YES
	observaciones		Varchar (100)	NO
FK	estado		Char (1)	YES

Figura 8 - Modelo de datos de actas

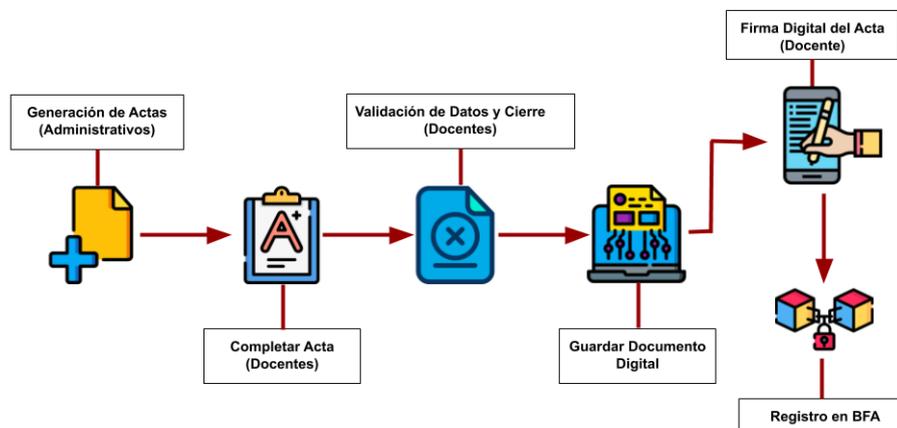


Figura 9 -Nuevo circuito de actas

Impacto sobre la comunidad universitaria

Desde los inicios del proyecto SIU-Guaraní, su implementación en las instituciones universitarias ha tenido un fuerte impacto en la calidad de los servicios brindados a alumnos, docentes, personal administrativo y autoridades. Ha sido, por ejemplo, el primer software del estado argentino en brindar servicios a través de teléfonos móviles, permitiendo a los alumnos realizar sus inscripciones por ese medio desde el año 2006.

Esta nueva funcionalidad no es la excepción, ya que transforma profundamente la manera en la que las instituciones llevan adelante la gestión de la información académica. Esta transformación se da al brindar un mecanismo que permite llevar adelante una gestión completamente digital de la historia académica de un alumno, eliminando cualquier registro en papel de la misma. Como todo cambio genera muchas resistencias, una de ellas está relacionada directamente con los temores respecto de la validez de la información cuando la misma no está acompañada de un documento físico (en papel) y con una firma holográfica que lo respalde. El hecho de abandonar los libros de actas, una práctica muy arraigada en las instituciones ya que se realiza desde los inicios de las mismas, implica un profundo cambio de paradigma. El lento despliegue que la firma digital tuvo en las universidades representa otra de las resistencias, aunque el advenimiento de la firma digital remota pareciera ser una alternativa viable para sortear este inconveniente.

Es indudable que esta solución le imprime agilidad y simplicidad a procesos históricamente complejos, como es el control de egreso de un alumno. Este proceso en particular requiere controlar una a una las actas en papel donde se registraron las notas de cada actuación académica de un alumno. Son tareas que, si realizan en forma manual, insumen un gran tiempo y esfuerzo debido a que es necesario que una persona identifique y valide la información del registro académico del alumno con el archivo físico que originó dicha información. Incluso muchas veces es necesario, en este

esquema en papel, trasladar copias físicas de los libros entre distintas áreas de la Universidad para realizar las validaciones correspondientes. Contar con las actas firmadas digitalmente simplifican la búsqueda, visualización y validación de la misma por las personas. Blockchain, por su parte, podría simplificar más aún el proceso mediante un proceso de un control automático por el sistema, y dejando la intervención de una persona para realizar una validación visual sólo ante alguna inconsistencia detectada entre la información actual de la base de datos y la enviada oportunamente a la red de bloques.

De manera complementaria, la implementación de soluciones de estas características requiere de políticas fuertes de backup y resguardo de la información, como así también de políticas que brinden seguridad física de los dispositivos involucrados, de modo de asegurar la disponibilidad y calidad de la información digital para la Institución.

Aprendizajes

Trabajar con nuevas tecnologías para generar herramientas que permitan brindar más y mejores servicios a las instituciones siempre deja lecciones. En primer lugar entendimos que aplicar cambios profundos en los procesos requiere tiempo de maduración, es muy probable que, de no haber realizado el proceso de manera gradual y progresiva, su implementación hubiese fracasado.

Los cambios organizacionales van asociados a cambios culturales de las personas que son parte de estas organizaciones, y estos cambios siempre se producen a un ritmo mucho menor que los tecnológicos. Por lo que es muy importante entender estos procesos de cambio cultural y estar atentos al momento en que la organización permite la implantación de estas nuevas lógicas.

Uno de los factores fundamentales del éxito de la digitalización de las gestiones es que SIU-Guaraní logró entender con claridad el ritmo de maduración de las instituciones, aprovechando cada uno de los puntos de inflexión para proponer cambios que, en los comienzos, fueron muy disruptivos pero con el correr del tiempo se aceptaron como parte de las reglas, permitiéndonos avanzar desde la emisión de certificaciones electrónicas, la digitalización de los requisitos de ingreso, hasta llegar finalmente a la implementación de firma digital y blockchain en las actas.

Agradecimientos

A las distintas personas que llevaron adelante la Coordinación General del SIU y a la Dirección del Consejo Interuniversitario Nacional por haber construido a lo largo de más de veinte años un espacio de trabajo colaborativo en la comunidad universitaria, fomentando el desarrollo de proyectos con un alto grado de innovación respetando las necesidades concretas de las Universidades Nacionales.

A los integrantes del equipo de desarrollo y testing del SIU-Guaraní, a los integrantes del área de Integración Funcional que participaron de este proceso; a todas las personas

que de una u otra manera colaboran día a para que el sistema de gestión académica se encuentre un paso más adelante y a las áreas transversales del SIU que facilitan las tareas cotidianas y agilizan la interacción entre quienes conformamos la Comunidad SIU.

Referencias

- [1] Consejo Interuniversitario Nacional "ESTATUTO DEL CONSEJO INTERUNIVERSITARIO NACIONAL (CIN)",
<http://www.cin.edu.ar/descargas/estatuto.pdf>
- [2] Ley 25506 Firma Digital,
<https://www.boletinoficial.gob.ar/detalleAviso/primera/7220862/20011214?busqueda=1>
- [3] Decreto 2628 - Boletín Oficial de la República Argentina,
<https://www.boletinoficial.gob.ar/detalleAviso/primera/7230494/20021220?busqueda=1>
- [4] Firma Digital Remota,
<https://www.argentina.gob.ar/modernizacion/firmadigital/firmadigitalremota>
- [5] Blockchain Federal Argentina, <https://bfa.ar/>
- [6] Registro Público de Graduados Universitarios,
<https://registrograduados.siu.edu.ar>

Uso de *smart contracts* para validación de certificados en una red *blockchain*

Miguel Montes, Romina Racca, Alejandro Biagetti, Solange Spitale,

Universidad Nacional de Córdoba, Ciudad Universitaria, Córdoba, Argentina
miguel.montes@unc.edu.ar, romina.racca@unc.edu.ar, abiagetti@unc.edu.ar,
solange.spitale@unc.edu.ar

Resumen: Las universidades, como instituciones de educación superior, otorgan títulos y expiden certificaciones que avalan las actividades realizadas y los niveles alcanzados por sus estudiantes. Si bien es habitual que la información contenida en estos certificados provenga de sistemas de registro informatizados, también es frecuente que por cuestiones legales exista un respaldo en papel del resultado de ciertos actos académicos. Este esquema de doble registro genera una necesidad de validar los certificados contrastándolos con la información en papel, lo cual representa una carga de trabajo considerable. Para resolver este problema, en la Universidad Nacional de Córdoba hemos implementado un esquema de registro de actas utilizando contratos inteligentes (*smart contracts*) desplegados en una red blockchain, más precisamente en la Blockchain Federal Argentina, una red permissionada que utiliza el modelo de consenso de *prueba de autoridad*. En este trabajo describimos los detalles del problema, las características de una red blockchain que pueden contribuir a su solución, en qué consiste la Blockchain Federal Argentina, la solución implementada en nuestra universidad y las lecciones aprendidas en este proyecto.

Palabras Clave: Blockchain, smart contracts, validación de certificados.

Eje temático: Smartcampus y tendencias tecnológicas emergentes en la Universidad

1 Introducción

En este trabajo proponemos un sistema de verificación de certificados de estudio mediante el uso de contratos inteligentes en una red blockchain. Para poder describir las características de una solución, primero debemos comprender el problema que se pretende resolver, por lo que comenzaremos por describir la realidad normativa de la Universidad Nacional de Córdoba y sus mecanismos de gestión académica, luego discutiremos las características de una red *blockchain* y los problemas que nos permite resolver, y finalmente presentaremos la solución propuesta.

2. Gestión académica en la Universidad Nacional de Córdoba

2.1 Introducción

La Universidad Nacional de Córdoba (UNC) es una universidad pública argentina. Es un organismo autónomo, que dicta sus propios estatutos y normas de gobierno. Su autonomía tiene rango constitucional, consagrada por la reforma de la Constitución Nacional del año 1994 [1].

Fundada en 1613, es la universidad más antigua del país, y con más de 120.000 alumnos, una de las más grandes [2].

Si bien su máxima autoridad es la Asamblea Universitaria, ésta sólo se reúne por convocatoria expresa, por lo que el gobierno es ejercido por el Consejo Superior y el Rector. El Consejo Superior está conformado por el Rector, los Decanos, y representantes de todos los claustros: docentes, no-docentes, egresados y estudiantes. Todas las autoridades, tanto unipersonales como colegiadas, son elegidas por los miembros de la comunidad universitaria mediante elección directa.

La universidad está conformada por quince facultades, cada una de ellas gobernada por un Consejo Directivo y un Decano [3].

A nivel central, el equipo de gestión del rectorado está organizado en Secretarías y Prosecretarías. Es así que la responsabilidad de la gestión de TIC corresponde a la Prosecretaría de Informática.

2.2 Marco normativo

Los procedimientos de registro y archivo de la actividad académica han sido establecidos en distintas ordenanzas y resoluciones del Consejo Superior. En el contexto de la enseñanza de grado, tiene especial relevancia la Ordenanza N° 7/2004 [4]. En esta ordenanza se establecen:

- La organización funcional de las áreas de enseñanza de las distintas unidades académicas.
- Los procedimientos básicos relacionados con el registro del desempeño de los alumnos de grado.

- Las normas de archivo de la documentación de desempeño académico.
- Las normas mínimas de administración y seguridad del sistema informático de desempeño académico.

En su artículo 6º, ordena:

“Establecer que las Actas de Examen, firmadas por los docentes responsables integrantes del tribunal examinador, sobre soporte de papel, tienen el carácter de documentos públicos y son las únicas fuentes originales de información para certificar el desempeño académico de los estudiantes de grado de la Universidad Nacional de Córdoba. Los profesores firmantes son los únicos responsables de la nota del alumno escrita en las Actas de Examen.”

Como puede apreciarse, si bien se reglamenta el uso del sistema informático, también se establece que la fuente última de verdad son las actas en papel.

En cuanto al sistema informático, esta ordenanza establece que cada unidad académica debe o bien asumir por sí misma su gestión, o delegarla en la Prosecretaría de Informática.

2.3 Sistema informático de gestión académica

La Universidad Nacional de Córdoba utiliza SIU Guaraní como sistema informático de gestión académica. SIU Guaraní es un sistema desarrollado por SIU (Sistema de Información Universitaria), un organismo que desarrolla distintos sistemas de gestión dirigidos a las universidades nacionales, y que depende del Consejo Interuniversitario Nacional.

Si bien inicialmente cada facultad administraba su propio sistema de gestión académica, con el correr de los años esa administración se fue centralizando en la Prosecretaría de Informática, mediante el proceso de delegación establecido en la ordenanza. Es decir, cada facultad es dueña de los datos, y determina los usuarios y roles necesarios, pero la administración de las bases de datos, actualizaciones, personalizaciones del sistema y resolución de problemas técnicos son responsabilidad del personal de la Prosecretaría.

Esta unificación en la administración ha tenido también su correlato en la unificación de las bases de datos. Hasta el año 2017, por motivos en parte técnicos y en parte organizacionales, cada unidad académica tenía su propia instancia de base de datos. En el año 2017 comenzamos la migración a la versión 3 de Guaraní, y en este proceso se han ido unificando todas las bases de datos en una base única gestionada a nivel central. Esa migración culminó en el primer semestre de 2019 para las carreras de grado, y se encuentra en proceso la migración de las carreras de posgrado.

2.4 Actas y certificados

Como se ha mencionado, los resultados de las evaluaciones se plasman en *actas*, como la que se muestra en la figura 1. Podemos ver que el acta contiene una serie de datos propios del acta en sí (fecha, asignatura, tribunal, etc.) y una serie de *renglones*, cada uno de los cuales corresponde a un alumno y contiene información acerca de él (nombre, número de documento, calificación, etc.). Denominaremos al conjunto de

datos comunes *cabecera*, independientemente de si en la versión impresa figuran en el encabezado o en el pie del acta.

Las actas son cargadas en el sistema Guaraní, impresas, firmadas de manera ológrafa por los docentes integrantes del tribunal examinador, y archivadas por la oficialía de cada unidad académica en *libros de actas*. Cuando se imprime la versión final del acta para ser firmada decimos que el acta ha sido *cerrada*, y ya no puede ser modificada. Cualquier error que se detecte a partir de ese momento sólo puede ser corregido mediante un *acta rectificativa*.

UNC		Universidad Nacional de Córdoba 589 - Facultad de Odontología				
ACTA DE EXAMEN						
		Libro: 00001	Acta: 13180	Hoja 01/01		
		LLAMADO: Llamado Jun-Jul/04 (1)		20/06/2004		
		CÁTEDRA - MESA: B				
(15-00019) Química General I						
NÚMERO	APELLIDO Y NOMBRE	DOCUMENTO	INGR.	COND.	NOTA	FIRMA
38456789	CARRERA, MARIA	DNI 38456789	2016	L	8 (ocho/00)	_____
31123456	QUIROZ, MERCEDES	DNI 31123456	2007	L	7 (siete/00)	_____
42123456	PEREZ, CARLA	DNI 42123456	2018	L	Ausente	_____
GARCIA, MARIA JULIA - MARTINEZ, DORA - AGUERO, NATALIA-						
Observaciones:						
Córdoba, ___/___/___-						
Certifico que la/s firma/s que ha/n sido puesta/s en la presente Acta pertenece/n a: _____						
3	1	2	0	2		
Inscriptos	Ausentes	Examinados	Reprobados	Aprobados		
			(0-3)	(4-10)		
					Libro/Acta: 0201800388	Hoja 01/01

Figura 1. Modelo de acta de examen

A requerimiento del alumno, el sistema también emite *certificados*. Un ejemplo de certificado se muestra en la figura 2. Al igual que un acta, también está compuesto por una cabecera y renglones, aunque en este caso los renglones corresponden con asignaturas o actividades realizadas.

Un certificado puede ser *provisorio* o *definitivo*. En el caso del certificado analítico definitivo, la unidad académica debe poseer un procedimiento de verificación de que su contenido se corresponda con lo que está registrado en las actas.

Este requerimiento de verificar en actas forma parte de un conjunto de procedimientos destinados a garantizar la integridad de las calificaciones obtenidas por los alumnos. Estos procedimientos implican una serie de controles cruzados en las que ninguna persona tiene de por sí la capacidad de alterar las calificaciones sin que esa alteración sea detectada por otros funcionarios que intervienen en el trámite. La comisión de un acto fraudulento requiere pues la colusión de varias personas

pertenecientes a distintos departamentos.

Si no existiera la verificación en actas cabría la posibilidad de que los administradores del sistema modifiquen las notas almacenadas en la base de datos, y que esas modificaciones resulten en certificados fraudulentos.

UNC		Universidad Nacional de Córdoba 589 - Facultad de Odontología						
CERTIFICADO ANALÍTICO FINAL								
Carrera: Odontología - Plan 2008								
Título: 5254 Odontólogo								
Apellido: QUIROZ								
Nombres: MERCEDES								
Documento: DNI 31123456								
Materia	Cred	Fecha Examen	Cd	Nota aplazo	Nota aprobado	Libro	Acta	Pg
15-00019 Química General I		20/06/2004	L		7.00 (Siete)	00001	13180	1
15-00050 Biología Celular		15/10/2018	L		7.00 (Siete)	02018	00772	1
15-00027 Física I		20/06/2005	R		6.50 (Seis c/50)	00015	14517	1
15-00035 Matemática I		20/06/2006	R		5.50 (Cinco c/50)	00016	00062	2
15-02302 Laboratorio I		26/07/2005	R	3.00 (Tres)		00015	15789	1
15-02302 Laboratorio I		16/12/2006	R		8.00 (Ocho)	00016	13578	3
Total de materias rendidas: 5				Promedio con aplazos: 6.17				
Total de materias aprobadas: 5				Promedio sin aplazos: 6.80				
Total de aplazos: 1				Total de Créditos: 0				
El promedio sólo incluye a las materias que son consideradas promediabiles.								
Fecha de ingreso: 02/12/2003								
Fecha de egreso: 23/02/2019								
CÓRDOBA, Rep. Argentina, VEINTIDÓS DE FEBRERO DE DOS MIL DIECINUEVE								

Figura 2. Certificado Analítico

2.5 El problema de la verificación en actas y posibles soluciones

La contrastación de los certificados con las respectivas actas conlleva un trabajo considerable. Para cada renglón del certificado es necesario buscar el acta correspondiente, y verificar que los datos de dicho renglón coincidan con los registrados en el acta, como se aprecia en la figura 3.

Si bien es posible disminuir el trabajo utilizando técnicas de muestreo, el problema subsiste: ¿De qué forma se puede garantizar que el resultado registrado en la base de datos no ha sido alterado?

Guaraní tiene una serie de características destinadas a minimizar el riesgo, en la forma de controles de acceso y registros de auditoría. Pero esto no es suficiente: los controles de acceso no protegen contra abusos por parte del administrador de bases de datos (DBA), o incluso de usuarios con perfiles de datos privilegiados; y los registros de auditoría permiten detectar cambios a posteriori, pero tampoco protegen contra abusos del DBA, que puede modificar estos registros.

Una buena solución a este problema debe garantizar a los responsables legales de las calificaciones que estas no han sido alteradas. Claramente se trata de un problema de

integridad de los datos.

Esta identificación del tipo de problema nos permite esbozar una primera solución razonable: el uso de *firma digital*.

UNIVERSIDAD NACIONAL DE CÓRDOBA		589 - Facultad de Odontología						
CERTIFICADO ANALÍTICO FINAL								
Carrera: Odontología - Plan 2008								
Título: 5254 Odontólogo								
Apellido: QUIROZ								
Nombres: MERCEDES								
Documento: DNI 31123456								
Materia	Cred	Fecha Examen	Cd	Nota aplazo	Nota aprobado	Libro	Acta	Pg
15-00019 Química General I		20/06/2004	L		7.00 (Siete)	00001	13180	1
15-00050 Biología Celular		15/10/2016	L		7.00 (Siete)	02018	00772	1
15-00027 Física I		20/06/2005	R		6.50 (Seis c/50)	00015	14517	1
15-00033 Matemática I		20/06/2006	R		5.50 (Cinco c/50)	00016	00062	2
15-02302 Laboratorio I		26/07/2005	R	3.00 (Tres)		00015	15789	1
15-02302 Laboratorio I		16/12/2006	R		8.00 (Ocho)	00016	13578	3
Total de materias rendidas: 5				Promedio con aplazos: 6.17				
Total de materias aprobadas: 5				Promedio sin aplazos: 6.80				
Total de aplazos: 1				Total de Créditos: 0				
El promedio sólo incluye a las materias que son consideradas promediables.								
Fecha de ingreso: 02/12/2003								
Fecha de egreso: 23/02/2019								
CÓRDOBA, Rep. Argentina, VEINTIDÓS DE FEBRERO DE DOS MIL DIECINUEVE								

UNIVERSIDAD NACIONAL DE CÓRDOBA		589 - Facultad de Odontología	
ACTA DE EXAMEN			
Libro: 00001		Acta: 13180	
LLAMADO: Llamado Jun-Jul/04 (1)		Hoja 01/01	
CÁTEDRA - MESA: B		20/06/2004	
(15-00019) Química General I			
NÚMERO	APELLIDO Y NOMBRE	DOCUMENTO	INGR. COND. NOTA FIRMA
38456789	CARRERA, MARIA	DNI 38456789	2016 L 8 (ocho/00)
31123456	QUIROZ, MERCEDES	DNI 31123456	2007 L 7 (siete/00)
42123456	PEREZ, CARLA	DNI 42123456	2016 L Ausente
GARCIA, MARIA JULIA - MARTINEZ, DORA - AGUERO, NATALIA-			
Observaciones:			
Córdoba, / / .			
Certifico que la/s firma/s que ha/hn sido puesta/s en la presente Acta pertenece/n a:			
3	1	2	0
Inscriptos	Ausentes	Examinados	Reprobados
		(0-3)	Aprobados (4-10)
		Libro/Acta: 0201800388 Hoja 01/01	

Figura 3. Contraste de certificados contra actas

2.5.1 Solución basada en firma digital

La firma digital [5,6] es un mecanismo de protección de la integridad de un documento u objeto digital que garantiza que dicho documento

1. ha sido producido (o al menos procesado) por el firmante
2. no ha sido alterado desde la firma
3. no puede ser repudiado por el firmante (el firmante no puede negar la validez de su firma).

En la República Argentina tiene por ley [7] el mismo valor probatorio que la firma autógrafa. La Ley 25.506 establece en su artículo 12:

“La exigencia legal de conservar documentos, registros o datos, también queda satisfecha con la conservación de los correspondientes documentos digitales firmados digitalmente”

En la Universidad Nacional de Córdoba ya se utiliza la firma digital en diversos contextos. Numerosos actos administrativos sólo se firman digitalmente (por ejemplo, las resoluciones rectorales).

Es decir, que existe el respaldo legal para reemplazar las actas en papel por actas digitales. Sin embargo, el mero reemplazo del acta en papel por un acta digital no nos resuelve el problema, ya que ahora deberíamos contrastar los certificados contra las actas digitales. Lo que necesitamos es un mecanismo que **nos garantice la integridad de la base de datos**.

Por lo tanto lo que se requiere es que se firmen los contenidos de la base de datos. Una posible forma de resolver el problema consistiría en:

1. Representar el acta en un formato procesable por una computadora (por ejemplo JSON o XML).

2. Presentar este documento digital para que se pueda validar que sus datos se corresponden con la realidad.
3. Permitir la firma digital de dicho documento y almacenar el resultado.
4. Representar los certificados en un formato similar.
5. Validar los certificados contra estas actas digitales, aceptando el resultado solo si la firma del documento sigue siendo válida.

Este mecanismo resuelve algunos de nuestros problemas. Por ejemplo, ya no es posible para el DBA modificar un acta sin que esa modificación sea detectada, porque no posee la clave privada necesaria para generar una firma válida.

Sin embargo, tiene varios inconvenientes:

Es complejo. La firma digital no es fácil de implementar en escala. Para mantener la seguridad de la clave privada es necesario utilizar *tokens criptográficos*, es decir, dispositivos de *hardware* que almacenan la clave en forma segura, y estos dispositivos suelen tener problemas de compatibilidad con determinados sistemas operativos y *browsers*. El manejo de claves no es trivial. La pérdida del PIN o contraseña implica la generación de un nuevo par de claves y la revocación del anterior.

No protege contra la eliminación de actas. Si bien puede alegarse que esto no es un problema de integridad sino de *disponibilidad*, esto no es cierto en todos los casos. La ausencia de una calificación desfavorable puede afectar un promedio, y eso ya es un problema de integridad si esa situación no es detectada.

Existe el problema de las actas rectificativas. Cuando se detecta un error en un acta que ya está cerrada, la solución no consiste en corregir el acta, sino en elaborar una nueva acta, denominada *acta rectificativa*, que salva el error. Pero si el acta original está firmada digitalmente, esa firma sigue siendo válida. No importa que se genere una nueva firma, si alguien intenta validar un certificado contra el acta original y la firma original, el certificado parecerá válido aunque ya no lo sea.

La validación debería ser hecha fuera del sistema. Si es el mismo sistema el que valida su propia salida, podría dudarse de su integridad. El sistema debe por lo tanto brindar a algún sistema externo las actas necesarias para validar. Pero esto a su vez es otro problema: estamos dando demasiada información. Para validar un certificado de un alumno estamos dando información sobre todos los alumnos que rindieron con él.

Ninguna de estas dificultades es insuperable, sin embargo nos obligan a recurrir a herramientas adicionales. Por ejemplo, deberíamos prever un mecanismo encadenado de firmas para que no se puedan borrar actas.

Cuando analizamos las características que debe tener la solución notamos que muchas de ellas coinciden con las provistas por una red *blockchain*. ¿Podría un mecanismo basado en blockchain resolver nuestro problema?

3. Blockchain

3.1 Generalidades

El término *blockchain* (cadena de bloques) está indeleblemente ligado al concepto de las llamadas criptomonedas, desde que fuera propuesto por Satoshi Nakamoto como

base de Bitcoin [8].

Bitcoin es un esquema de dinero electrónico basado en una red "peer-to-peer", en la cual todos los participantes comparten un único registro distribuido de transacciones. Para asegurar la inmutabilidad del registro, las transacciones se almacenan en bloques ligados criptográficamente entre sí, de manera que es imposible modificar un bloque sin modificar todos los bloques posteriores. Nuevas transacciones se registran en nuevos bloques que se agregan a la cadena utilizando un *mecanismo de consenso* que garantiza que todos los participantes de la red reconozcan como válida la misma versión de la cadena. Si todos los participantes acuerdan en la validez de la misma secuencia de transacciones, entonces todos están también de acuerdo en cuanto al estado que deriva de la aplicación de esas transacciones.

Con posterioridad a Bitcoin surgieron otras redes: Ethereum, Monero, ZCash son algunos ejemplos. Todas comparten el mismo concepto de cadena de bloques ligados criptográficamente, aunque pueden diferir en los algoritmos utilizados y en el mecanismo de consenso.

Algunas de estas redes son abiertas, es decir, cualquiera puede participar instalando un nodo y conectándolo a la red. Otras redes requieren de algún tipo de autorización para participar. Estas últimas se denominan *permisionadas*.

Entre los mecanismos de consenso utilizados vale la pena mencionar los siguientes:

Proof-of-Work (PoW): Es el esquema original utilizado en Bitcoin. Básicamente consiste en una lotería, en el cual el nodo de la red que encuentra una solución a cierto problema gana el derecho a agregar un nuevo bloque. Para encontrar la solución hay que realizar cierto trabajo computacional de costo relativamente reducido, pero con bajísima probabilidad de éxito. Por lo tanto es necesario repetirlo muchas veces hasta encontrar la solución. En el caso de Bitcoin, este problema consiste en calcular un doble hash de un bloque mediante el algoritmo SHA-256, de manera que el resultado sea inferior a cierto número prefijado. El nivel de dificultad de esta búsqueda se ajusta para que dados todos los recursos computacionales aplicados a la tarea en un determinado momento, el tiempo entre bloques sea aproximadamente 10 minutos. Este sistema es simple en el sentido en que cualquiera puede participar, pero implica un costo considerable en energía eléctrica y tiene un impacto ambiental para nada despreciable.

Proof-of-Stake (PoS): El problema del alto costo de PoW condujo al desarrollo de otros mecanismos, como el de prueba de interés (Proof-of-Stake). En este caso la capacidad para agregar nuevos bloques deriva de la inversión que el nodo tenga en la cadena. La red Ethereum, por ejemplo, usa PoW pero planea migrar en un futuro a PoS.

Proof-of-Authority (PoA): En este esquema sólo un conjunto de nodos autorizados, llamados *selladores*, puede agregar bloques a la cadena. Este esquema es típico de las redes permisionadas, y requiere la existencia de un algoritmo que determine el orden en que operan los selladores.

3.2 Ethereum

A partir de la creación de Bitcoin surgieron propuestas para utilizar las características de una red blockchain en la solución de otros problemas, además del de implementar dinero electrónico.

Reviste particular interés el caso de Ethereum, basada en una propuesta de Vitalik

Buterin [9]. La idea es implementar en la red una *máquina virtual descentralizada* que pueda ejecutar código arbitrario. Es decir, las transacciones almacenadas en los bloques no representan únicamente transferencias de fondos, como en el caso de Bitcoin, sino que especifican acciones a ejecutar en esa máquina virtual. Como el mecanismo de consenso permite que todos los nodos de la red tengan una misma vista de la cadena de bloques, todos ejecutan la misma secuencia de instrucciones de la máquina virtual y todos llegan a un mismo estado.

Claramente no es un mecanismo eficiente para realizar una computación, ya que esta se repite en múltiples nodos. Pero es útil si el problema requiere precisamente que todos puedan realizarla y verificarla.

Los programas que se ejecutan en esta máquina virtual (EVM, *Ethereum Virtual Machine*)[10] se denominan *smart contracts* (contratos inteligentes), y el lenguaje preferido para desarrollarlos se llama Solidity [11]. Para desplegar un contrato en la red

1. se escribe el contrato en Solidity
2. se compila a *bytecode* de la EVM
3. se despliega en la red mediante una *transacción*.

La interacción con el contrato se realiza mediante *transacciones* y *llamadas*. Una transacción sirve para invocar funciones del contrato que modifican el estado de la cadena, mientras que una llamada (*call*) es una consulta que obtiene información pero no modifica la cadena.

Ethereum tiene, además, su propia criptomoneda, llamada *ether*, que puede utilizarse y transferirse de manera similar a Bitcoin, y que tiene su propia valuación de mercado. Pero además el ether sirve como forma de pago de las computaciones realizadas. Cada instrucción ejecutada en la EVM, así como el almacenamiento utilizado tiene un costo en *gas*. Podríamos decir que el gas es el combustible que hace funcionar la EVM. El gas necesario para las distintas operaciones está prefijado, pero ese gas debe pagarse con ether a un costo determinado por leyes de mercado. Esto proporciona un sistema de incentivos económicos para el correcto funcionamiento de la red.

Las aplicaciones que corren en la EVM se denominan *Dapps* (*decentralized applications*). Su arquitectura implica como mínimo un *backend* que se ejecuta en la red, y un *frontend* que suele ser una aplicación web. El *frontend* se puede desarrollar en distintos lenguajes y utilizando distintos *frameworks*. Por ejemplo, una de las opciones más comunes es hacerlo en lenguaje *Javascript* utilizando la biblioteca *web3.js* [12].

3.3 Tipos de problemas para los que es aplicable una red blockchain

Blockchain es una herramienta, y como toda herramienta debe aplicarse sólo a los problemas para los cuales es adecuada. Muchos problemas que pretenden resolverse con una red blockchain se pueden resolver de manera mucho más eficiente con una solución centralizada.

La clave aquí es la descentralización. ¿Requiere mi problema una solución descentralizada? Si es así, tal vez vale la pena analizar una solución basada en blockchain.

En ese sentido es muy útil un documento elaborado por NIST (National Institute of Standards and Technology) [13], que incluye un diagrama de flujo destinado a ayudar en la decisión sobre si es adecuado usar blockchain para un determinado problema. Este

diagrama, presentado en la figura 4, fue elaborado originalmente por The United States Department of Homeland Security (DHS) Science & Technology Directorate, y reconoce como válidos los casos de uso en los que:

1. Se requiere un almacén de datos compartido y consistente.
2. Existe más de una entidad que aporta datos o se trata de un caso de auditoría.
3. Los registros escritos no serán nunca borrados ni modificados.
4. No se almacenarán datos sensitivos.
5. Es difícil determinar quién debe controlar el almacén de datos.
6. Se requiere una bitácora (log) de todo lo registrado.

3.4 La Blockchain Federal Argentina (BFA)

La Blockchain Federal Argentina es una plataforma multiservicios, gobernada siguiendo el modelo de *Múltiples Partes Interesadas* (multistakeholder). Participan organizaciones gubernamentales, de la academia, empresas y organizaciones de la sociedad civil. La Universidad Nacional de Córdoba ha participado en la BFA desde su comienzo, y es uno de los cuatro selladores que comenzaron a operar la red el 27 de septiembre de 2018.

Desde el punto de vista técnico opera una red basada en Ethereum, con Proof-of-Authority como mecanismo de consenso. Al momento de escribir este trabajo cuenta con 23 selladores, número planificado desde su inicio, aunque se está analizando la posibilidad de aumentarlo.

La participación es gratuita. Si bien se utilizan los conceptos de *gas* y *ether* propios de Ethereum, se provee a los miembros del *ether* necesario sin costo.

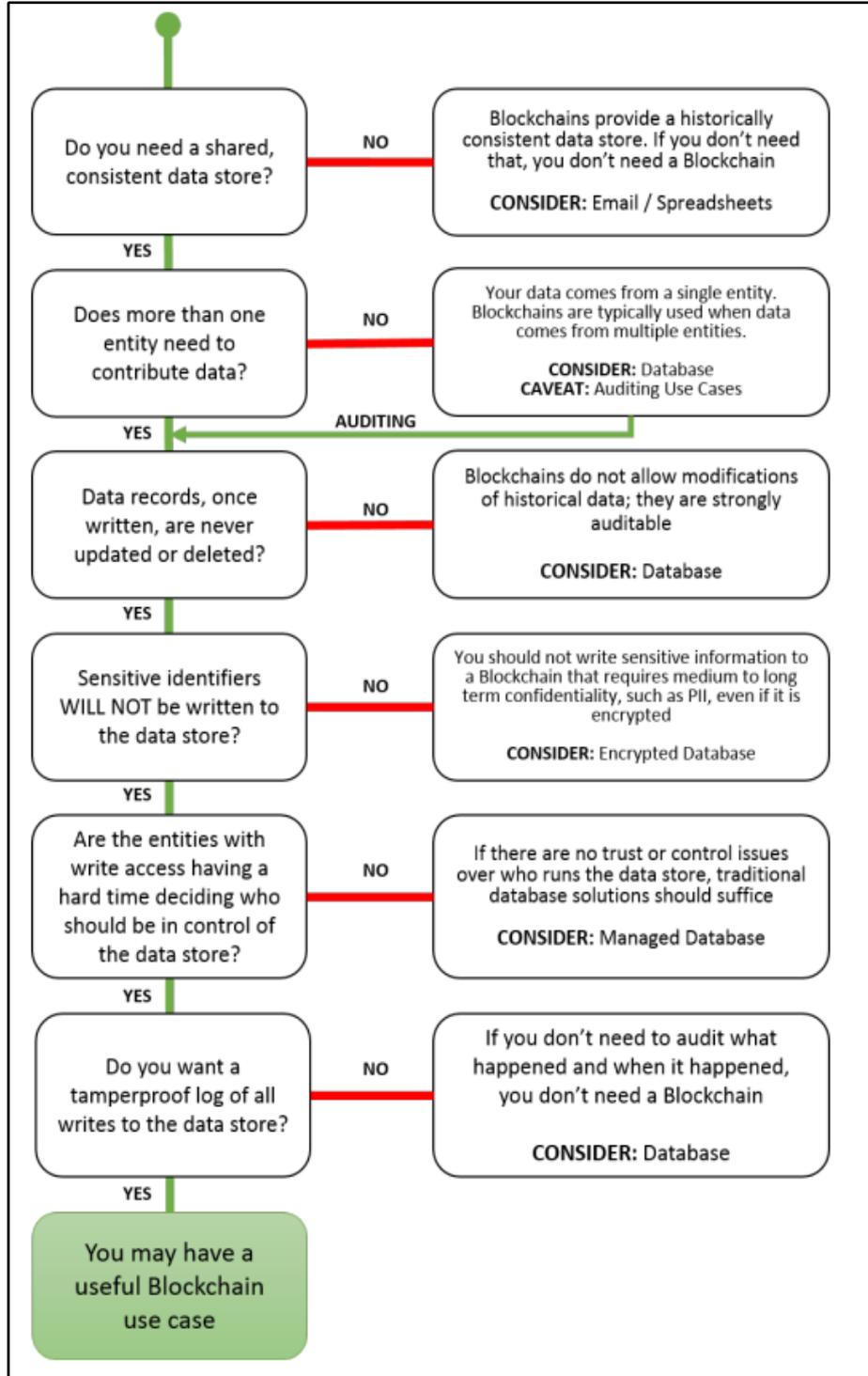


Figura 4. Diagrama de flujo para evaluar la pertinencia de una solución blockchain

4. La solución implementada en la UNC

Como mencionamos más arriba, la firma digital es una solución parcial a nuestro problema, ya que no resuelve problemas tales como la eliminación de actas, o los asociados con las actas rectificativas. Notamos que lo que se requiere es un mecanismo que no permita modificar los datos una vez que se han registrado, y la blockchain parece darnos una herramienta adecuada para implementar ese mecanismo.

Analicemos el problema con la secuencia de preguntas propuestas por NIST:

¿Necesitamos un almacén de datos compartido y consistente? Sí, queremos tener un registro de las actas que sea verificable por múltiples partes interesadas.

¿Existe más de una entidad que aporta datos o se trata de un caso de auditoría? Los datos provienen de una fuente principal, que es el sistema Guaraní. Sin embargo, queremos que estos datos puedan ser auditados por múltiples partes.

¿Los registros escritos nunca serán borrados ni modificados? Ese es justamente el objetivo buscado por el sistema.

¿No se almacenarán datos sensibles? El diseño del sistema deberá tener en cuenta que así sea.

¿Es difícil determinar quién debe controlar el almacén de datos? Precisamente deseamos que *nadie* tenga control absoluto del almacén de datos.

¿Se requiere una bitácora de todo lo registrado? Sí, queremos que exista un registro de todas las acciones realizadas.

4.1 Descripción de la solución

El objetivo es registrar en la blockchain una *evidencia* de las actas emitidas por el sistema. Decimos evidencia en el sentido de que lo que se guarda no es el acta misma (no debemos guardar datos sensibles en la cadena), sino una prueba de que una cierta acta existía en un determinado momento.

Para ello modelamos el acta como un *smart contract*, en el cual se guarda un resumen o digesto criptográfico de la cabecera y de cada uno de los renglones. Ese resumen se obtiene aplicando una *función de hash criptográfica*, y tiene la propiedad de que es computacionalmente imposible encontrar otro mensaje o documento que produzca el mismo resumen.

Para validar un certificado lo recorremos renglón por renglón, identificamos el acta que corresponde a cada renglón, y verificamos que el resumen construido a partir del certificado coincide con el resumen registrado en el contrato.

El sistema busca no tener falsos positivos, pero puede tener falsos negativos. Es decir, si el sistema dice que un certificado ha sido verificado, es porque se corresponde con las actas. Por el contrario, si dice que no ha podido verificarlo simplemente dice que con la evidencia existente en la cadena no puede certificar que es válido, y debe validarse por otro medio (por ejemplo, contrastando contra las actas originales). La figura 5 muestra una vista de los componentes del sistema.

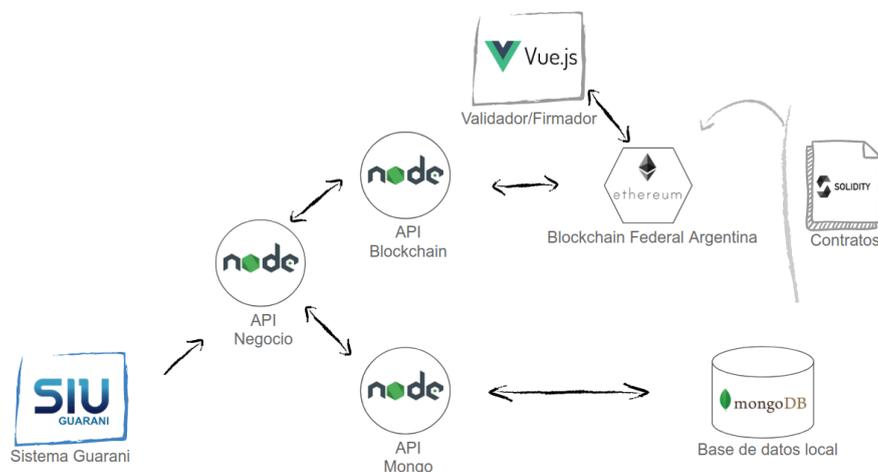


Figura 5. Componentes del sistema

Contratos desplegados en la red. Una versión preliminar del sistema de contratos desplegados ha sido descrito previamente [14]. Existen:

- Actas: Cada acta individual está modelada en un contrato, que provee los métodos necesarios para la validación de renglones. Las actas pueden ser firmadas en forma independiente por un número arbitrario de personas.
- Actas rectificativas: Las actas rectificativas son modeladas por otro contrato que tiene en cuenta sus particularidades, y registran a qué acta rectifican.
- Registro de actas: Para que no se deba lidiar con miles de direcciones de contratos individuales, existe un registro de actas que permite asociar actas con organismos emisores y encontrar el acta deseada.
- Contrato de roles: Asigna roles a usuarios para determinar quienes pueden registrar actas en representación de una determinada institución.

Cabe destacar que se ha previsto que los contratos puedan ser utilizados por cualquier institución que tenga un modelo de actas y certificados similar al planteado en este trabajo.

Cierre y registro de actas. El sistema Guarani cierra las actas y se comunica con una API REST transmitiendo el contenido del acta cerrada en formato JSON. Por un lado se registra dicha acta en la cadena, y por otro lado se preserva en una base de datos MongoDB.

La razón para preservar el acta es que cuando se emite un certificado es necesario poder reconstruir la realidad existente al momento del cierre de las actas.

El acta contiene muchos datos que pueden cambiar con el paso del tiempo. Por ejemplo, el alumno puede cambiar de nombre o de número de documento. Ambas son realidades que no están modeladas adecuadamente en el sistema de gestión académica. Es decir, el sistema registra el nombre y el número de documento actual, pero no es capaz de determinar qué nombre y número de documentos estaban registrados en el momento de un determinado examen.

Esta necesidad de preservar el acta en una base separada puede desaparecer si, en alguna versión futura de Guarani, eso comienza a ser provisto por el sistema.

Firma de actas. Una vez que las actas están registradas en la Blockchain, los funcionarios responsables de su validez pueden firmarlas. Esa firma queda registrada en el contrato, y se informa a quienes posteriormente validan certificados. Un ejemplo de la vista que recibe un firmante se aprecia en la figura 6.

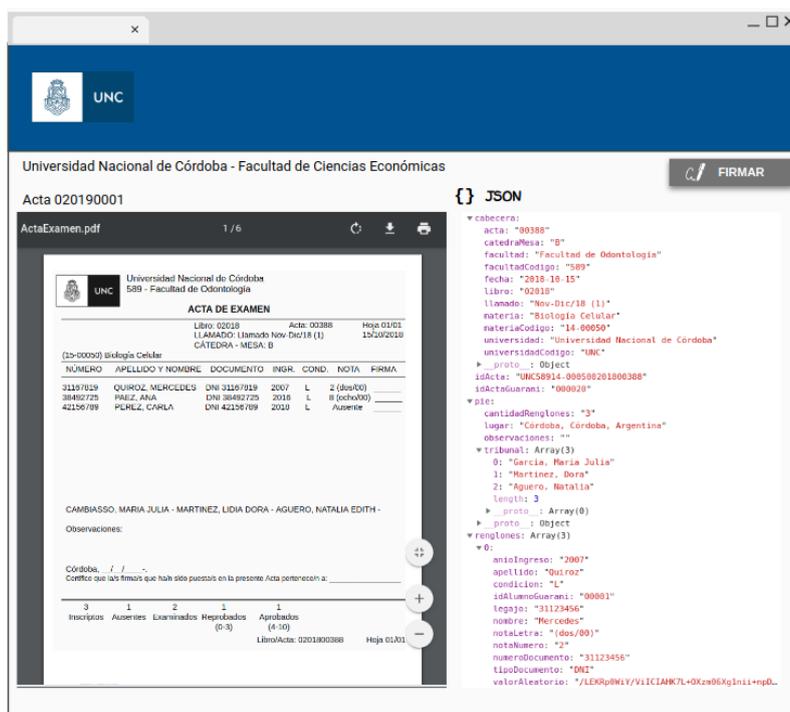


Figura 6. Firma de acta en Blockchain

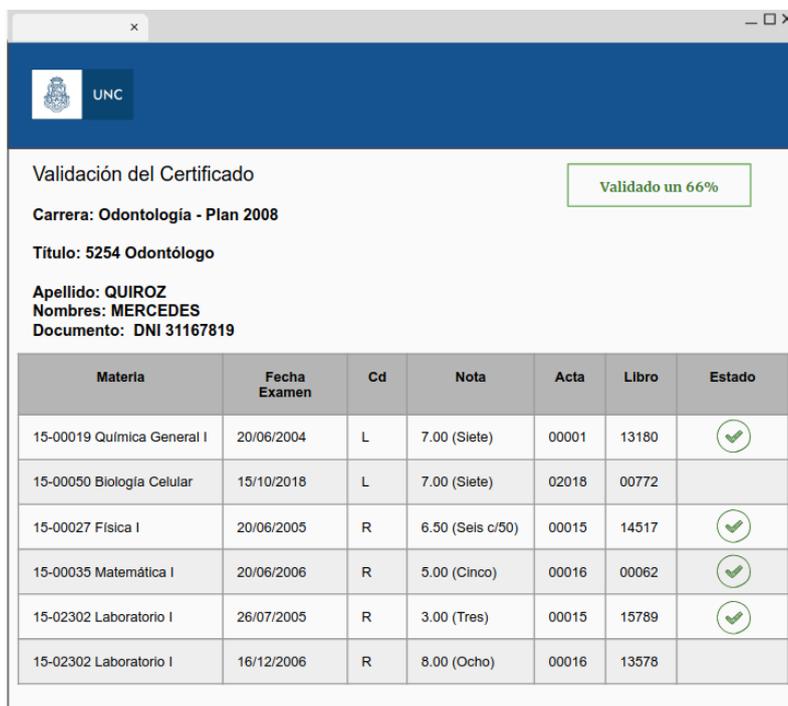
Emisión de certificados. El sistema Guaraní emite certificados PDF que contienen incrustado un certificado en formato JSON que es validable. Para ello usa su propia información y la almacenada en la base de datos MongoDB.

Validación de certificados. Se ha desarrollado un validador externo que recibe un certificado validable (con JSON incrustado), y verifica los resultados contra las actas registradas en la cadena. Un ejemplo de validación de certificado se muestra en la figura 7.

Es necesario que ni la firma de las actas ni la validación de certificados requieran utilizar las herramientas desarrolladas por la UNC. Ambas operaciones se realizan interactuando directamente con la cadena a través de cualquier nodo transaccional de la BFA. El código de los contratos desplegados es público, por lo que cualquiera puede desarrollar su propia versión del firmador o del validador.

El objetivo de este diseño es maximizar la transparencia. Se registran las actas en el momento del cierre, y a partir de ese momento la universidad no tiene posibilidades de modificar lo que ha declarado. Existe un registro auditable de todas las acciones realizadas. Y al mismo tiempo, dado que no se guardan datos primarios en la cadena,

no se compromete ningún tipo de información sensible. No es posible obtener de la cadena nombres de alumnos, números de documento ni calificaciones.



Validation of Certificate

Validado un 66%

Carrera: Odontología - Plan 2008

Título: 5254 Odontólogo

Apellido: QUIROZ
Nombres: MERCEDES
Documento: DNI 31167819

Materia	Fecha Examen	Cd	Nota	Acta	Libro	Estado
15-00019 Química General I	20/06/2004	L	7.00 (Siete)	00001	13180	✓
15-00050 Biología Celular	15/10/2018	L	7.00 (Siete)	02018	00772	
15-00027 Física I	20/06/2005	R	6.50 (Seis c/50)	00015	14517	✓
15-00035 Matemática I	20/06/2006	R	5.00 (Cinco)	00016	00062	✓
15-02302 Laboratorio I	26/07/2005	R	3.00 (Tres)	00015	15789	✓
15-02302 Laboratorio I	16/12/2006	R	8.00 (Ocho)	00016	13578	

Figura 7. Validación de certificado

5. Lecciones aprendidas y trabajo futuro

El desarrollo de esta aplicación ha brindado una serie de lecciones interesantes en cuanto existen diferencias fundamentales con el desarrollo de aplicaciones tradicionales.

El código desplegado en la blockchain no puede ser modificado. Es necesario pensar detenidamente lo que esto implica. En una aplicación tradicional, si se detecta un *bug*, se corrige y se despliega la versión corregida. Los contratos desplegados, por el contrario, viven por siempre. Entre las medidas que es necesario aplicar podemos mencionar:

- Existencia desde un comienzo de un mecanismo de versionado, que permita incorporar funcionalidades mediante nuevas versiones de un contrato.
- Máxima simplicidad de los contratos. Los contratos deben ser lo más sencillos posibles, y delegar funcionalidades no esenciales en programas externos.

No se puede modelar la solución basándose en los modelos transaccionales existentes. Existe la tentación de basarse en modelos existentes (en nuestro caso, Guarani). Pero las diferencias de enfoque son tan grandes que eso no es posible. Hay

que analizar el problema que queremos resolver y modelarlo con una nueva visión.

Los contratos deben ser simples. Lo hemos mencionado antes, pero esta necesidad va más allá de la imposibilidad de modificación. Los contratos deben ser fáciles de comprender, y su semántica debe ser clara simplemente leyendo el código.

Nuestros sistemas transaccionales registran insuficiente información histórica. Como hemos mencionado, nuestros sistemas carecen de la capacidad de obtener de manera sencilla el estado en un cierto momento del pasado. Si bien todos los cambios se registran, esto se hace con fines de auditoría. Y durante mucho tiempo hemos considerado ciertos atributos de una persona como inamovibles (por ejemplo el género), cuando la realidad nos ha probado que no lo son. Esto dificulta el desarrollo de sistemas como este, en los cuales necesitamos poder reconstruir el estado en un cierto momento (como por ejemplo un examen).

Como trabajo futuro, debemos analizar la posibilidad de modificar nuestros sistemas transaccionales para que registren la información necesaria de manera nativa. Esto reduciría la necesidad de guardar información duplicada en una base de datos adicional.

También debemos trabajar fuertemente en los aspectos de usabilidad. Si bien hemos tratado de simplificar al máximo la operación por parte de usuarios no técnicos, el manejo de claves sigue siendo un tema conflictivo que es necesario mejorar.

En el estado actual del proyecto no proponemos la eliminación de las actas en papel, pero otra línea de trabajo futuro es avanzar en ese sentido. Una vez afianzado este mecanismo de registro, podremos analizar los cambios en la normativa necesarios para el reemplazo de las actas en papel por algún otro sistema de resguardo.

6. Conclusiones

Consideramos que este sistema es un paso en la dirección correcta de asegurar la transparencia de nuestros sistemas de registro, y que es un caso de uso adecuado para la herramienta seleccionada, blockchain.

El sistema ha sido modelado de la forma más genérica posible, y hemos previsto su uso por distintas instituciones educativas. Puede ser utilizado por cualquier institución que registre sus instancias de aprobación utilizando actas, y que emita certificados analíticos basados en ellas. En particular, este modelo es aplicable a todas las universidades argentinas, porque refleja la normativa existente a nivel nacional.

También interpretamos que es un caso de uso más que valida la existencia de la BFA, en conjunto con los demás casos ya implementados o en desarrollo.

Como hemos mostrado en nuestro análisis, las soluciones basadas en blockchain son adecuadas en un estricto conjunto de casos, pero creemos que nuestro proyecto reúne todas las características necesarias.

8. Referencias

1. Constitución de la Nación Argentina, (1994)
<http://servicios.infoleg.gob.ar/infolegInternet/anexos/0-4999/804/norma.htm>
2. Arnaldo Mangeaud et al. Anuario Estadístico 2017. Universidad Nacional de Córdoba.

- ISBN 978-987-778-889-1 (2018).
3. Estatuto de la Universidad Nacional de Córdoba <https://www.unc.edu.ar/sobre-la-unc/estatuto>
 4. Ordenanza HCS 7 (2004) http://www.digesto.unc.edu.ar/consejo-superior/honorable-consejo-superior/ordenanza/7_2004
 5. Alfred J. Menezes, Paul C. van Oorschot y Scott A. Vanstone. "Handbook of Applied Cryptography". CRC Press (1997)
 6. Jean-Philippe Aumasson. "Serious Cryptography". No Starch Press Inc. (2019)
 7. Ley 25.506 - Ley de Firma Digital (2001)
<http://servicios.infoleg.gob.ar/infolegInternet/anexos/70000-74999/70749/norma.htm>
 8. Satoshi Nakamoto. "Bitcoin: A Peer-to-Peer Electronic Cash System" (2006)
 9. Vitalik Buterin. "Ethereum White Paper. A next generation smart contract & decentralized application platform". (2013)
 10. Gavin Wood. "Ethereum: A Secure Decentralised Generalised Transaction Ledger". (2018)
 11. "Solidity 0.5.8 documentation". <https://solidity.readthedocs.io/en/v0.5.8/>
 12. "web3.js - Ethereum Javascript API" <https://web3js.readthedocs.io/en/1.0/>
 13. Dylan Yaga, Peter Mell, Nik Roby y Karen Scarfone. "Blockchain Technology Overview" NISTIR 8202. <https://doi.org/10.6028/NIST.IR.8202> (2018)
 14. Miguel Montes. "Verificación de actas y certificados usando una red Blockchain". 6º Congreso Nacional de Ingeniería en Informática/Sistemas de Información. ISSN 2347-0372. (2018)

Ambiente integral para fortalecer la seguridad de la información en los sistemas de la Universidad de Guadalajara

Luis Alberto Gutiérrez Díaz de León, Sergio Javier Uribe Nava, Jesús David Salas Valle, Miriam Elizabeth Huerta Raygoza.

Universidad de Guadalajara, Av. Juárez 976,
CP 44100, Guadalajara, Jalisco.
miriam.huerta@redudg.udg.mx, david.salas@redudg.udg.mx

Resumen. Este documento se inscribe en el eje temático de Smartcampus y tendencias tecnológicas emergentes en la Universidad (aplicaciones de la inteligencia artificial, Blockchain, analítica, IoT, impresión 3D, experiencia inmersiva), el cual busca aumentar la seguridad de la información, proteger la privacidad de la misma y obtener una evolución significativa en la implementación y adopción de las tecnologías de información. Su objetivo es proveer una mejora en el servicio de transacciones electrónicas, así como la confiabilidad y seguridad de las mismas. El proyecto impacta a la Red Universitaria en la mejora de la competitividad y seguridad de la información de los sistemas de la institución. El primer sistema que usó la Red Privada Blockchain RPB es el Sistema de Entrega-Recepción Universitario SERU, este permite llevar el control y seguimiento del proceso de registro de información de las Entregas-Recepción de los cambios de puestos en la administración de la UDG. El problema que atiende el proyecto es que no existía un mecanismo de protección para la integridad de los expedientes electrónicos los cuales requieren de la trazabilidad en sus transacciones, registros únicos y no alterables. Los usuarios afectados por la práctica resultante son 2916 con 4356 expedientes vinculados con la RPB. El trabajo realizado consiste en que una vez cerrado el expediente se propaga en la RPB donde se identifica mediante un consenso con los nodos si el documento es íntegro, de ser así se confirma la validación.

El equipo de trabajo fue multidisciplinario conformado por 36 personas, su alcance es estatal y tiene una proyección nacional.

Palabras Clave: Blockchain, cadena de bloque, contrato inteligente, sistema, tecnología, transacción, criptografía, algoritmo, plataforma de desarrollo, red descentralizada, criptomoneda, seguridad, trazabilidad, comunidad SNMP, monitor, usuario y expediente.

Eje temático: Smartcampus y tendencias tecnológicas emergentes en la Universidad (aplicaciones de la inteligencia artificial, Blockchain, analítica, IoT, impresión 3D, experiencia inmersiva).

1. Introducción

Con el presente proyecto se busca utilizar eficientemente la Red Privada Blockchain de la Universidad de Guadalajara, la cual fue realizada a principios del año 2018 y se presentó en la Octava Conferencia de Directores de Tecnología de Información, TICAL 2018. Cabe mencionar que en esta nueva etapa se sumaron nuevas tecnologías y sistemas para el uso de la misma, es decir, se incorporó un monitor para la Red Privada Blockchain -RPB- como componente adicional, el cual consiste en monitorear los nodos participantes. Esta hecho con tecnologías diferentes a las que se utilizan en la Coordinación General de Tecnologías de Información -CGTI- y sus características principales son: la integración, el rendimiento, la seguridad, la sincronización y la tolerancia a fallas. Por otro lado, se implementó la comunidad Simple Network Management Protocol –SNMP- junto la herramienta PRTG Network, dicha fusión consiste en la creación de una comunidad específica para la RPB que permite monitorear aproximadamente más de 100 sensores de la red en los nodos participantes, está configurada con los parámetros de dominio, IPV4, IPV6 y el tipo de servidor – virtual y/o físico-. Es a través de la comunidad SNMP con la herramienta PRTG Network que se determina la salud de los nodos participantes.

Finalmente, a dichas implementaciones se sumó la incorporación del Sistema de Entrega-Recepción Universitario –SERU- con el objetivo de seguir la trazabilidad en sus transacciones, registros únicos y no alterables.

2. Problemática y contexto en que se propuso el proyecto

La problemática y contexto en que se propuso la incorporación de SERU a la RPB es debido a que la comunidad universitaria y la sociedad exige al sector público un rendimiento claro y puntal de cuentas.

La Universidad de Guadalajara, como universidad pública del Estado de Jalisco, debe inscribirse dentro de los avances y transformaciones que hoy en día experimentan los órganos de fiscalización en el país. Por lo que se requiere contar con instrumentos más efectivos y eficaces de rendición de cuentas, transparencia y fiscalización que garanticen el uso honesto y eficiente de los recursos que la sociedad aporta a la Universidad de Guadalajara, así como, un oportuno y claro rendimiento de cuentas por parte de los funcionarios universitarios. De acuerdo a lo establecido en H. Consejo General Universitario, a propuesta del Rector General, aprobó el Reglamento del Sistema de Fiscalización de la Universidad de Guadalajara (RSF), con fecha 15 de marzo de 2002, Contraloría General de conformidad con el ordenamiento antes citado, tiene como atribución establecer los lineamientos para la entrega-recepción de las dependencias universitarias, de conformidad con la fracción XX del artículo 48 del RSF, con apoyo al sistema de fiscalización y con el fin de uniformar e incorporar al Sistema de Entrega-Recepción Universitario –SERU-, la información sobre los recursos humanos, financieros, materiales, de control escolar y archivos de las dependencias y garantizar la continuidad en la prestación de los servicios universitarios, de los proyectos y programas en un proceso de entrega-recepción [1].

Una vez mencionado lo anterior, se procede a explicar que, aunque el proceso de las

Entregas-Recepciones se lleva a cabo a través del sistema SERU, existe de manera intermitente la incertidumbre por parte de los usuarios acerca de la seguridad de la información y/o la veracidad de la misma, debido a que no existía un mecanismo de protección para la integridad de los expedientes electrónicos los cuales requieren de la trazabilidad en sus transacciones, registros únicos y no alterables. Adicionalmente, surgió de la necesidad de reducir sustancialmente el uso de papel físico.

3. Descripción de la solución tecnológica implementada

Una vez mencionado el problema al que atiende la incorporación de SERU a la RPB y considerando que dicho trabajo fue realizado sobre la experiencia implementada anteriormente, se procede a describir brevemente la misma.

La RPB permite garantizar que la información de cualquier sistema que haga uso de la misma sea íntegra, logrando así detectar si tuvo o no modificaciones por algún usuario. Su característica principal es la inmutabilidad, certeza y transparencia para la Universidad de Guadalajara.

Adicional a lo anterior, una de las principales características de la RPB es la personalización, debido a que el proyecto consistió en incorporar la tecnología Blockchain para fortalecer la plataforma de desarrollo de software de la CGTI, y así ofrecer un servicio robusto y seguro para el desarrollo de sistemas que cumplan las necesidades de los procesos de la Universidad de Guadalajara.

Para finalizar, la mejora de sistemas también forma parte de las principales características del proyecto, debido a que la RPB permite realizar transacciones no financieras, confiables y de manera segura, agiliza la administración de los actores en la red, evita la manipulación de información y reduce la duplicidad de trabajo en el desarrollo de los sistemas de información.

Los procesos que se pueden implementar en la RPB son:

- Transacciones que requieren trazabilidad y *exigen registros únicos y no alterables*,
- Procesos de identidad, y
- Nuevos modelos de negocio.

La solución tecnológica que se implementó para subsanar el mecanismo de protección para la integridad de los expedientes electrónicos fue incorporar SERU a la RPB, contemplando la seguridad de la información en los expedientes cerrados y gestionados en la Red Universitaria para el proceso de Entrega-Recepción, el complemento es a través de la incorporación de un ambiente integral de la plataforma Blockchain institucional, en el momento que un expediente es cerrado por contraloría general, es decir concluyó el proceso de Entrega-Recepción de una dependencia, este expediente se convierte en una transacción para la RPB, la cual provee mecanismos para poder vincular y validar las transacciones de los procesos sistematizados de la institución. Dicho proceso se muestra a continuación gráficamente:

Relación SERU-Blockchain

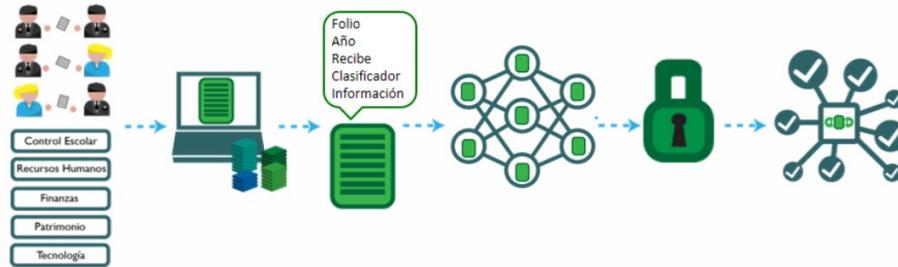


Fig 1. Relación SERU-Blockchain

A continuación, se muestra la evidencia de cómo está construida la RPB y dicha incorporación:

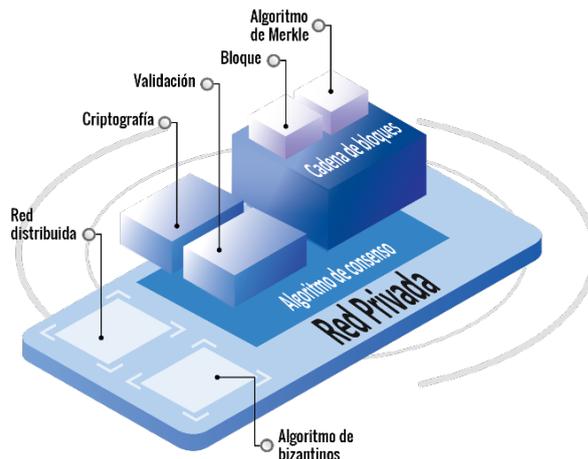


Fig. 2 Componentes de la Red Privada Blockchain.

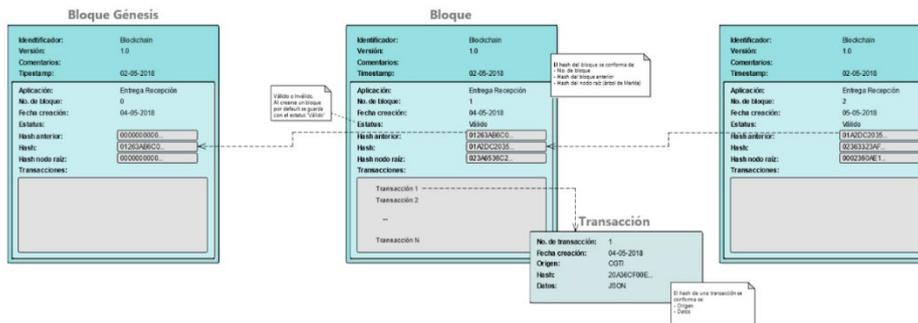


Fig. 3 Cadena de bloques.



Fig. 4 Cadena de bloques desde el aplicativo web (1).

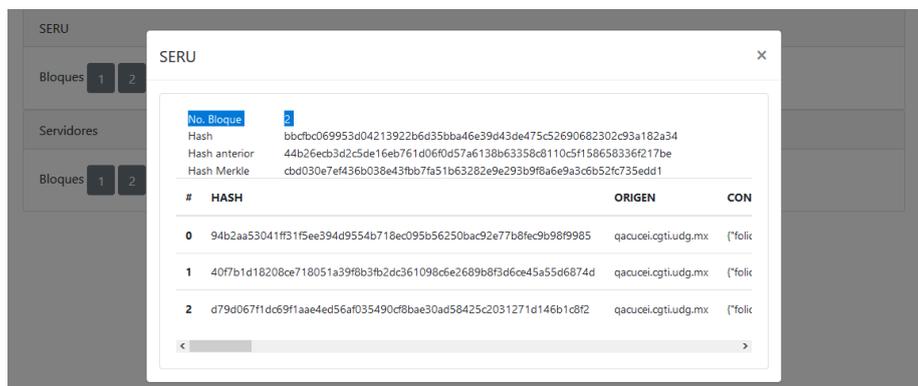


Fig. 5 Cadena de bloques desde el aplicativo web (2).

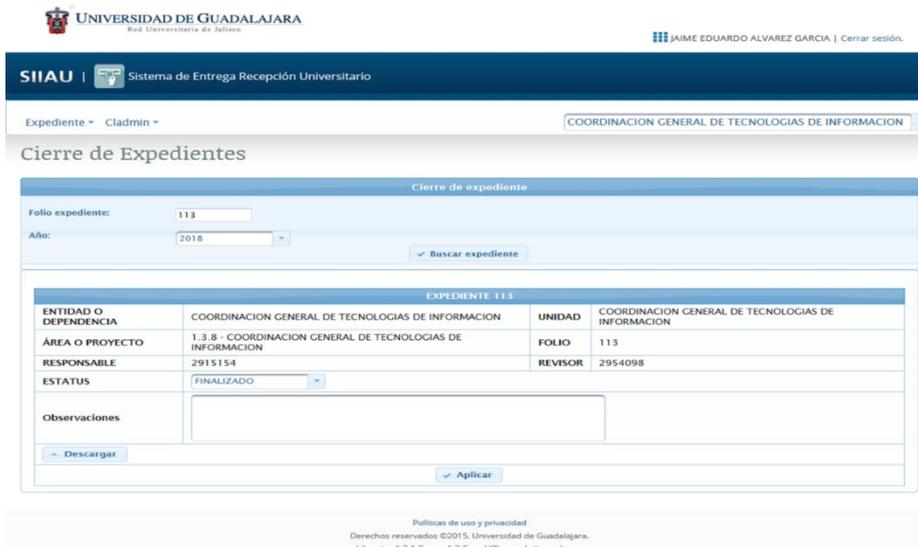


Fig. 6 Cierre de un expediente.

UNIVERSIDAD DE GUADALAJARA
Red Universitaria de Jalisco

JAIME EDUARDO ALVAREZ GARCIA | Cerrar sesión.

SIIAU | Sistema de Entrega Recepción Universitario

Expediente - Cladmin - COORDINACION GENERAL DE TECNOLOGIAS DE INFORMACION

Consulta de Expedientes

Expediente

Rol: REVISOR
Folio expediente: 113
Año: 2018
Buscar

Encabezado

UNIVERSIDAD DE GUADALAJARA		FECHA:
ENTIDAD O DEPENDENCIA:	COORDINACION GENERAL DE TECNOLOGIAS DE INFORMACION	COORDINACION GENERAL DE TECNOLOGIAS DE INFORMACION
UNIDAD:	COORDINACION GENERAL DE TECNOLOGIAS DE INFORMACION	
ÁREA O PROYECTO:	1.3.8 - COORDINACION GENERAL DE TECNOLOGIAS DE INFOR	ESTATUS EXPEDIENTE: CERRADO
Clasificación:	--Selección--	
Formato:	--Selección--	

No hay registros por mostrar (1 de 1) 10

Expediente-PDF's Validación Blockchain

Fig. 7 Consulta de un expediente.

UNIVERSIDAD DE GUADALAJARA
Red Universitaria de Jalisco

JAIME EDUARDO ALVAREZ GARCIA | Cerrar sesión.

SIIAU | Sistema de Entrega Recepción Universitario

Expediente - Cladmin - COORDINACION GENERAL DE TECNOLOGIAS DE INFORMACION

Consulta de Expedientes

Expediente

Rol: REVISOR
Folio expediente: 113
Año: 2018
Buscar

Encabezado

UNIVERSIDAD DE GUADALAJARA		FECHA:
ENTIDAD O DEPENDENCIA:	COORDINACION GENERAL DE TECNOLOGIAS DE INFORMACION	COORDINACION GENERAL DE TECNOLOGIAS DE INFORMACION
UNIDAD:	COORDINACION GENERAL DE TECNOLOGIAS DE INFORMACION	
ÁREA O PROYECTO:	1.3.8 - COORDINACION GENERAL DE TECNOLOGIAS DE INFOR	ESTATUS EXPEDIENTE: CERRADO
Clasificación:	--Selección--	
Formato:	--Selección--	

No hay registros por mostrar (1 de 1) 10

Expediente-PDF's Validación Blockchain

Éxito
El contenido del expediente es válido.

Fig. 8 Caso de éxito.

Expediente | Cladmin | COORDINACION GENERAL DE TECNOLOGIAS DE INFORMACION

Consulta de Expedientes

Expediente

Rol: REVISOR
Folio expediente: 113
Año: 2018
[Buscar]

Encabezado

UNIVERSIDAD DE GUADALAJARA | FECHA: 11/12/2018

ENTIDAD O DEPENDENCIA: COORDINACION GENERAL DE TECNOLOGIAS DE INFORMACION | UNIDAD: COORDINACION GENERAL DE TECNOLOGIAS DE INFORMACION
ÁREA O PROYECTO: 1.3.8 - COORDINACION GENERAL DE TECNOLOGIAS DE INFOR | ESTATUS EXPEDIENTE: CERRADO

Clasificación: --Selección--
Formato: --Selección--

No hay registros por mostrar
(1 de 1) | 10

Expediente-PDF's | Validación Blockchain

ENTREGÓ:	2915154	[Buscar código]	RECIBÍÓ:	2915154	[Buscar código]
NOMBRE:	JESUS DAVID SALAS VALLE		NOMBRE:	JESUS DAVID SALAS VALLE	
CARGO:	JEFE		CARGO:	JEFE	
OBSERVACIONES:			ESTATUS:		

Fig. 9 Caso de error.

Un componente adicional que está incorporado a la RPB es el monitor el cual consiste en monitorear los nodos participantes, esta creado con tecnologías diferentes a las que se utilizan en la CGTI y sus características principales son: la integración, el rendimiento, seguridad, sincronización y la tolerancia a fallas. El monitor permite determinar a través una interfaz que nodos participantes se encuentran activos y sincronizados en terminos de información y/o red, así como la validación de transacciones.

Las herramientas con las que esta creado el monitor son Phyton 3 como lenguaje de programación para la lógica del sistema, Falcon framework para la creación de la API REST, PostgreSQL como gestor de base de datos y React para crear la interfaz de usuarii web con el lenguaje de programación JavaScript. Para verificar la funcionalidad del monitor se realizaron pruebas funcionales individuales para cada proceso y al final de forma conjunta.

A continuación, se muestra la evidencia del monitor

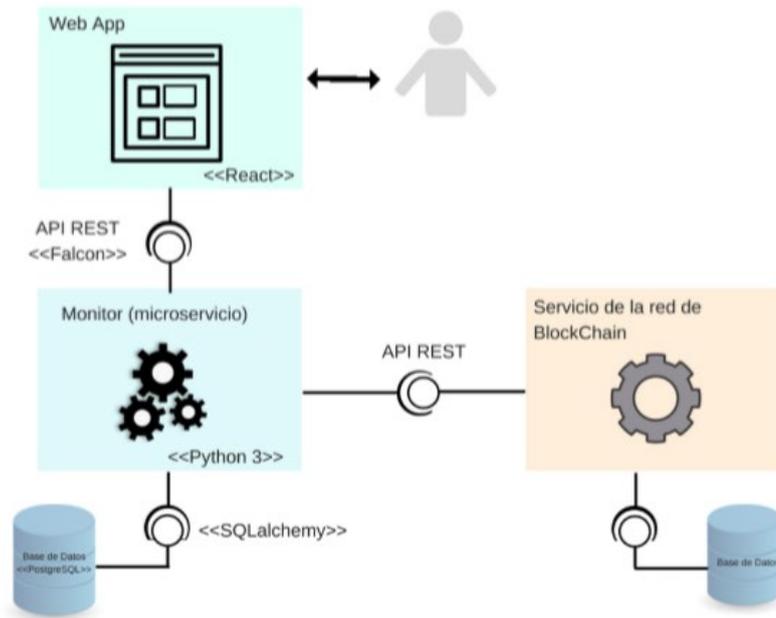


Fig. 10 Diagrama de arquitectura de software

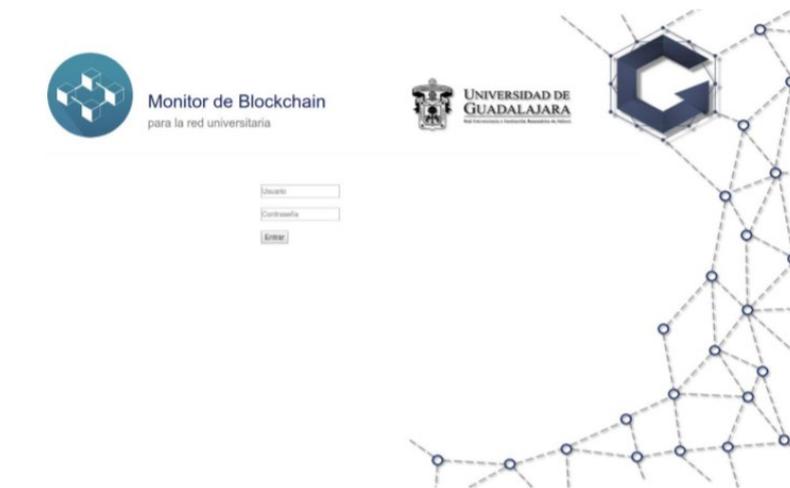


Fig. 11 Inicio y cierre de sesión.

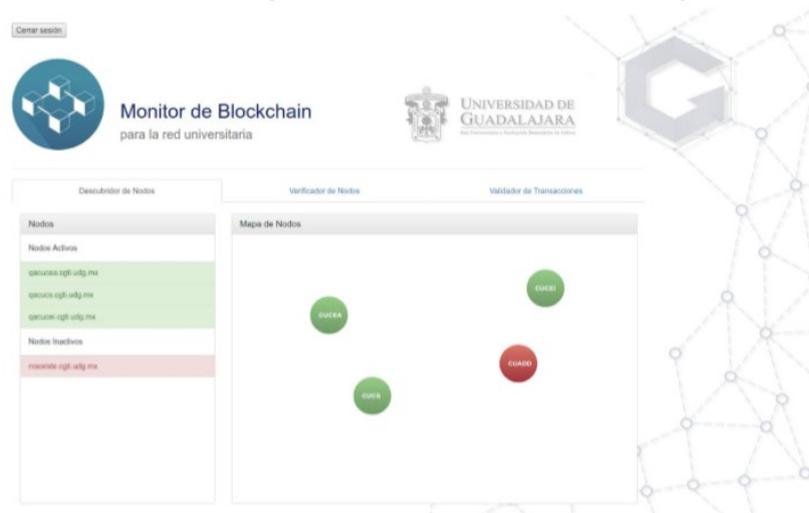


Fig. 12 Descubridor de nodos.

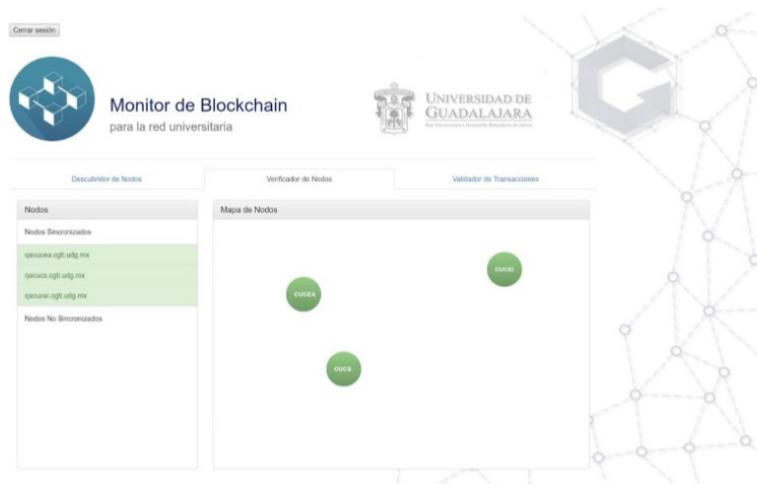


Fig. 13 Verificador de nodos.

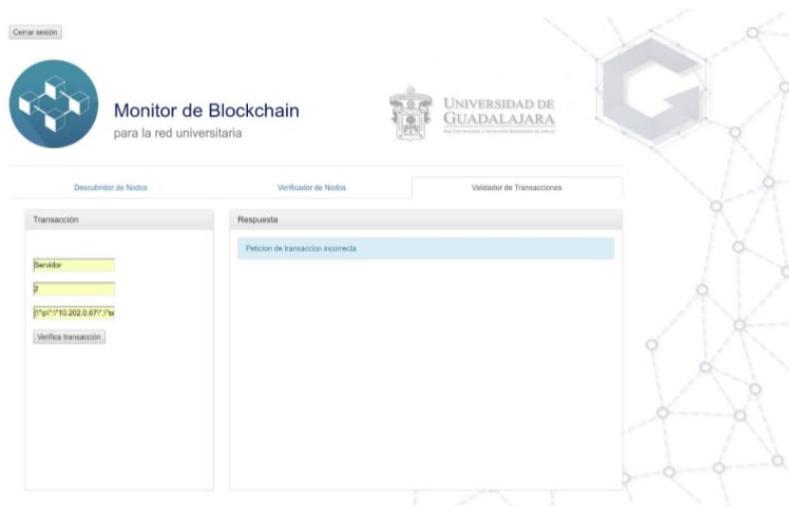


Fig. 14 Verificador de transacciones.

El último componente adicional que está incorporado a la RPB es implementación de la comunidad SNMP con la herramienta PRTG Network, dicha implementación consiste en una comunidad específica para la RPB y permite monitorear 100 sensores de red de los nodos participantes, está configurada con los parámetros de dominio, IPV4, IPV6 y el tipo de servidor –virtual y/o físico-.

La comunidad SNMP es un protocolo de la capa de aplicación que facilita el intercambio de información de administración entre dispositivos de red y permite a los administradores supervisar el funcionamiento de la red, buscar y resolver sus problemas, y planear su crecimiento. A continuación, se muestra la evidencia de la comunidad SNMP con la herramienta PRTG Network:

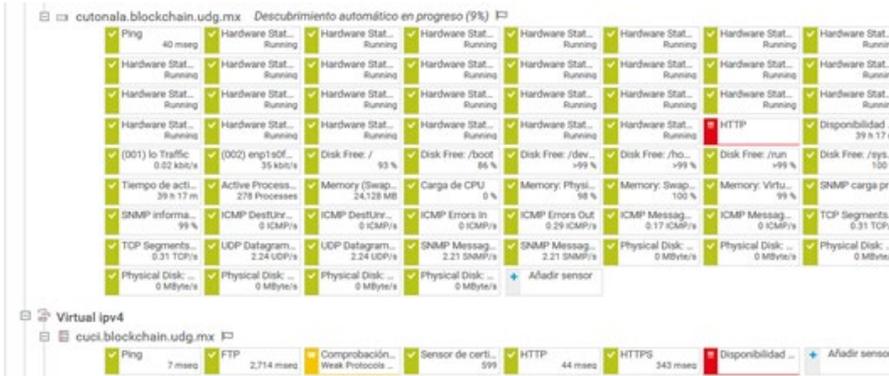


Fig. 15 Monitoreo de algunos nodos a través de la comunidad SNMP.

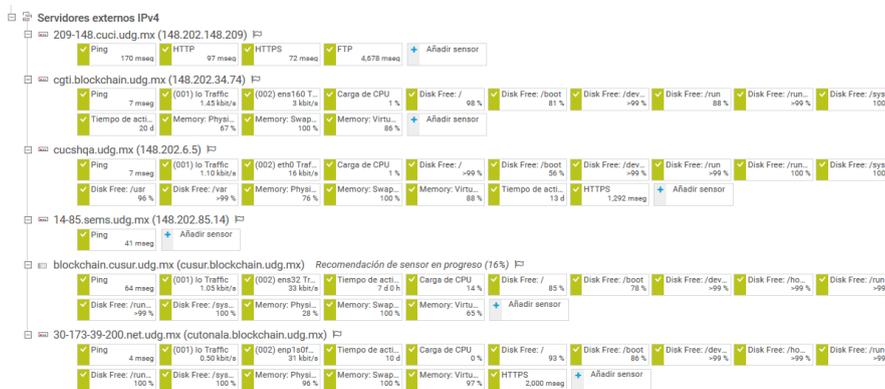


Fig. 16 Monitoreo del nodo CUTONALA a través de la comunidad SNMP.

4. Aspectos críticos y relevantes a resaltar y detallar

La Universidad de Guadalajara cuenta con una cantidad extensa de sistemas diseñados,

desarrollados e implementados sobre una plataforma de desarrollo propia de la CGTI, la cual permite cumplir diversas necesidades de la institución, debido a que se tienen sistemas y/o aplicaciones hechos a la medida. Una vez mencionado lo anterior, uno de los aspectos relevantes a resaltar es que se optó por seguir la misma directriz, es decir, diseñar, desarrollar e implementar una plataforma de tecnología Blockchain segura y hecha a la medida para la red universitaria, así como un monitor para la Red Privada Blockchain y la creación de la comunidad SNMP con la herramienta PRTG Network.

El segundo aspecto relevante a resaltar es que para la ejecución del trabajo se integró un equipo de trabajo multidisciplinario, donde participaron compañeros de las diferentes áreas de la Universidad de Guadalajara, las cuales corresponden a la CGTI, Maestría en Tecnologías de Información los CTAs de los Centros Universitarios de CUSUR, CUTTONALA, CUCI y CUCSH, así como la Coordinación de Cómputo e Informática del Sistema de Educación Media Superior.

Por otra parte, también se integraron personas voluntarias que resultó de su interés colaborar en conjunto, como, por ejemplo: prestadores de servicio y becarios de la Ingeniería en informática, Computación e Industrial. Los 36 colaboradores fueron un elemento muy importante en el desarrollo e implementación del trabajo debido a la división de roles y actividades para la obtención de un mejor resultado.

Cabe mencionar que el tercer aspecto relevante a resaltar es el haber utilizado la infraestructura física o virtual existente en los diferentes centros universitarios y en la Coordinación de Cómputo e Informática de la Universidad de Guadalajara, esto para lograr una red distribuida y como consecuencia el costo de inversión en el proyecto fue mínima.

El cuarto aspecto relevante a resaltar fue la creación del monitor de la RPB y la implementación de la comunidad SNMP con la herramienta PRTG Network, debido a que son componentes adicionales de la RPB que proveen un beneficio de suma importancia ya que permiten monitorear los nodos en el aspecto de la información, validar transacciones y en cuestiones de red, determinar la salud de los nodos participantes. Cabe mencionar que tanto el monitor y la comunidad SNMP con la herramienta PRTG Network fueron implementados por alumnos voluntarios de la Maestría en Tecnologías de Información y formaron parte del equipo multidisciplinario.

El quinto aspecto relevante a resaltar es la innovación tecnológica al ser una Universidad pública de América Latina que cuenta con una RPB utilizando un desarrollo propio y hecho a la medida.

5. Resultados obtenidos y su impacto

Como se mencionó anteriormente, el primer sistema que hizo uso de la RPB en conjunto con el monitor y la comunidad SNMP es SERU, el cual consiste en llevar el control y seguimiento del proceso de registro de información de las Entregas-recepción que se generan a partir de los cambios de puestos de la Administración de la Universidad de Guadalajara.

Hasta el momento los usuarios afectados por la práctica resultante son:



Fig. 17 Usuarios afectados por la práctica resultante.

El presente trabajo está impactando a gran parte de la comunidad universitaria por su influencia transversal en la seguridad de la información y en la innovación tecnológica, ya que permite obtener un crecimiento significativo en la adopción de las TICS mediante el desarrollo propio y completamente seguro.

6. Aprendizajes

Tabla 1. Aprendizajes

Universidad de Guadalajara		
Coordinación General de Tecnologías de Información		
Aprendizajes		
	¿Qué Trabajo Bien?	¿Qué Puede Ser Mejorado?
Definición de Requerimientos y Administración	<p>Se trabajó de manera cercana con los interesados y equipo de trabajo para obtener los requerimientos y la información técnica fácilmente.</p> <p>Es importante considerar la figura del líder del proyecto debido a que él va guiando la ruta del proyecto y proporciona un gran apoyo en los temas técnicos que el administrador del proyecto desconoce. La definición de los requerimientos también se obtiene con el cliente y los procuradores.</p>	<p>La colaboración de algunas personas importantes para el avance programático del proyecto se percibió un poco dispersa debido a que con frecuencia no se concretaban las actividades en tiempo y forma, ya que por lo regular estaban ocupados con otras actividades prioritarias.</p>

<p>Definición de Alcance y Administración</p>	<p>Se cumplieron en forma los requerimientos principales que los interesados definieron desde un inicio en el proyecto. Incluso se añadieron factores de seguridad, monitoreo de redes y persistencia de la base de datos para un mejor funcionamiento de dicha red, así como los nodos participantes.</p>	<p>Se considera que al inicio del proyecto no se dimensiono de la forma correcta los requerimientos ni el tiempo en el que se desarrollarían. Se había dimensionado que la Red Privada Blockchain y la Plataforma para Contratos Inteligentes se terminarían en 6 meses cada uno. Sin embargo, nos tomó un año en realizar la Red Privada Blockchain y la Plataforma de Contratos Inteligentes se comenzaría a partir de marzo 2019. El equipo de trabajo del proyecto considera que esto se debe a que en nuestra coordinación se estaban atendiendo diversos proyectos con mayor prioridad, por lo tanto, tenían que dedicarles más tiempo a otras actividades.</p>
<p>Costo Estimado y Control</p>	<p>Se considera que el único costo que hubo en el proyecto fue el recurso humano y una parte en el mantenimiento de los servidores asignados a pruebas y uno a producción. Sin embargo todos los nodos participantes fueron prestados voluntariamente por cada CTA.</p>	<p>Reducir el tiempo de la ejecución del proyecto, de tal manera que no se puedan extender mucho tiempo porque en ese caso el costo del recurso humano si se eleva en una gran parte, debido a que ellos podrían atender otras actividades.</p>
<p>Planeamiento de Calidad y Control</p>	<p>No existe una política sobre la calidad mínima que debe de tener el producto desarrollado, sin embargo, se debe de entregar con errores mínimos y la mayor parte de requerimientos implementados. Como parte de la calidad, en la CGTI existe un área de pruebas funcionales las cuales aseguran un nivel mínimo de calidad.</p>	<p>Definir las políticas de calidad, respecto al desarrollo del producto, así como el nivel de calidad esperado en la atención de los usuarios.</p>
<p>Disponibilidad de Recursos Humanos, Equipos de Desarrollo y Desempeño</p>	<p>Uno de los aspectos más importantes para la elaboración del proyecto es la administración de las actividades, tiempos y recurso humano. Por lo que en el caso de este proyecto se realizó una planeación para llevar a buen término el desarrollo del mismo, sin embargo, hubo cambios o modificaciones que corrompían completamente el</p>	<p>Generar una asignación del personal con un compromiso de trabajo, no basta con solo firmar el documento y mencionar que está asignado, si no dedicar tiempo a las actividades y cumplir con las tareas asignadas. De lo contrario</p>

	<p>cronograma de actividades y el plan de trabajo. Esto se debía a que el personal en diversas ocasiones estaba ocupado en otras actividades o a las decisiones de los directivos.</p> <p>A pesar de la prioridad e importancia de los proyectos, si los intereses del recurso humano no están alineados con los intereses del proyecto, se perderá mucho tiempo y dinero.</p>	<p>cuando el personal no tiene la disponibilidad requerida es conveniente cerrar el proyecto por falta de apoyo.</p>
Administración de Comunicación	<p>Se realizó un plan para la gestión de comunicaciones, el cual involucra a los interesados, el cliente y el equipo de trabajo.</p>	<p>A pesar de los canales de comunicación establecidos, el equipo de trabajo no los respetó ni prestó atención a la información que se les proporcionaba, aun así, argumentaban que no se les informaba.</p>
Administración de Grupos de Interés	<p>Mantener informados a los interesados es un factor muy importante para el proyecto, ya que ha permitido que el trabajo se realice a pesar de los problemas que se han presentado y además siempre están informados con los correos de informe de avance periódicos.</p>	-
Reportajes	<p>Aproximadamente entre 15 días y un mes se les enviaba un correo informativo a los directivos de esta coordinación, en estos se consideraban fechas propuestas, actividades realizadas, porcentaje de avance, actividades pendientes, problemas presentados, posibles soluciones y una proyección de fechas de cierre.</p>	<p>Un factor a mejorar podría ser la frecuencia con la que se enviaban los correos, sin embargo, esta se definió así debido a la prioridad del proyecto.</p>

Referencias

[1] Universidad de Guadalajara. (2016). Circular 01/2016 (oficio No. 0598/2016). Recuperado de <http://www.contraloriageneral.udg.mx/sites/default/files/16%20Contraloria%20Gral.%20Oficio%20No.%200598-2016%2017-marzo-16.pdf>

Capítulo 6

Respondiendo a la universidad cambiante de la era digital mediante desarrollos rápidos y flexibles

Sistema integral para la implementación de una credencial universitaria

Luis Alberto Gutiérrez Díaz de León, Sergio Javier Uribe Nava, Jesús David Salas Valle, Ana Nayeli Briones Islas, David Armando Rios Cervantes.

Universidad de Guadalajara, Av. Juárez 976,
CP 44100, Guadalajara, Jalisco.
nayeli.briones@redudg.udg.mx , david.salas@redudg.udg.mx

Resumen. Este documento se inscribe en el eje Respondiendo a la universidad cámbiate de la era digital mediante desarrollos rápidos y flexibles, el cual busca generar una estrategia para la implementación del cambio en las credenciales de identificación como miembro de la Universidad de Guadalajara, que actualmente son otorgadas a los alumnos activos y que se encuentran vigentes. El proyecto impacta a toda la matrícula de la Institución poniendo a su disposición dos sistemas integrados que le permiten registrar su solicitud, generar el cargo correspondiente, dar seguimiento a su trámite, recibir notificaciones, entre otras funcionalidades, hasta la impresión y entrega de su credencial universitaria, la cual está fabricada con material de PVC compuesto y contiene un chip de identificación por radiofrecuencia, que la hacen susceptible a distintos servicios y aplicativos tales como, control de acceso a instalaciones, control de asistencia en las aulas, préstamos y control de libros en las bibliotecas universitaria, interacción con el Sistema de Transporte Público del Estado de Jalisco.

El equipo de trabajo fue multidisciplinario conformado por 20 personas, su alcance es estatal beneficiando a una matrícula de 288,000 estudiantes aproximadamente.

Palabras Clave: Credencial, tarjeta inteligente, tecnología de identificación por radiofrecuencia, usuario, alumno, SIIAU, proximidad, Mifare Desfire, Unique ID.

Eje temático: Respondiendo a la universidad cámbiate de la era digital mediante desarrollos rápidos y flexibles.

1. Introducción

Con el presente proyecto se busca otorgar una credencial universitaria inteligente a cada alumno de la Institución, la cual le permite acceder a distintos servicios internos y externos a la Universidad, el cambio de la credencial tiene como finalidad unificar el formato e información que presentan las credenciales universitarias, y que además cuenten con las mismas características tecnológicas. Cabe mencionar que el nuevo diseño de la credencial se encuentra apegado al nuevo marco legal de protección de datos personales, cuenta con niveles de seguridad y facilita la identificación visual y digital. Para ello se realizaron grandes actividades como: la adquisición de equipamiento para la personalización de las credenciales y consumibles necesarios, se renovó el diseño de la credencial universitaria, desarrollo de un sistema que permite llevar el seguimiento de los trámites, impresión, vinculación de UniqueID (Identificador Único) de cada credencial con el código de alumno asignado, y la gestión de los diferentes estatus. Por otro lado, los directivos mantuvieron comunicación con la Secretaría de Transporte Público del Estado.

La integración del Sistema de Credencialización Universitaria fortalece al SIAU como un sistema más para maximizar los beneficios que esto conlleva, tales como mantener comunicación y vinculación con los diferentes procesos que se requieren para validar la información.

2. Problemática y contexto en que se propuso el proyecto

La problemática y contexto en que se propuso la implementación del proyecto y el cambio, es debido a que las credenciales otorgadas contaban con el diseño y formato lanzado hace 13 años aproximadamente, el cual solo permitían identificación visual, no disponían de mecanismos de seguridad para validar su autenticidad o bien con tecnologías para vincularse con otros servicios.

Otro aspecto a considerar fue que los alumnos debían contar con múltiples credenciales para diferentes servicios como: acceso a la escuela o instalaciones universitarias como laboratorios y bibliotecas, identificación oficial como alumno de la Universidad, recibir beneficios como becas o descuentos en entidades externas y transporte público, uso de transporte eléctrico público, entre otros. Es importante mencionar que cerca del 70% de las escuelas preparatorias de la Institución ya cuentan con la infraestructura para el servicio de control de acceso en sus instalaciones, de tal manera que la credencial viene a reforzar esta estrategia de seguridad en los campus.

Debido a lo anterior, y atendiendo al Plan de Desarrollo Institucional 2014-2030 en su eje estratégico "Gestión y Gobierno", el proyecto se encuentra alineado en los objetivos 15 y 16 con sus respectivas estrategias:

"Fortalecimiento de la gestión y gobernanza universitarias"

- Adoptar procesos de gestión mediante esquemas de gobierno electrónico para agilizar la toma de decisiones
- Implementar acciones para reducir las asimetrías y rezagos entre las entidades de la Red, respetando la diversidad y ventajas competitivas regionales.

“Sustentabilidad con transparencia y rendición de cuentas”

- Implementar una política para el ahorro y eficiencia del gasto universitario y mejorar la difusión sobre el uso de los recursos y su impacto en las funciones sustantivas.

3. Descripción de la solución tecnológica implementada

Una vez detallado el problema al que atiende la implementación del proyecto y considerando que dicho trabajo fue realizado en un tiempo aproximado de 18 meses y con la participación de 30 personas de diferentes áreas operativas y administrativas de la Institución, así como externas a ella, se procede a describir brevemente el proceso de implementación.

A partir de la generación del diseño de la nueva credencial universitaria estableciendo los datos de impresión y características visuales, se realiza la adquisición de equipamiento para la personalización de las mismas, inicia el levantamiento de requerimientos y reglas de operación del nuevo Sistema de Credencialización Universitaria y la adecuación al Sistema de Ventanilla Única de Servicios, definidas por la Coordinación General de Control Escolar -CGCE- quien tiene el rol de Cliente y las atribuciones de credenciar a los alumnos de la Universidad, según la normatividad.

Paralelo a estos trabajos se sostuvieron varias reuniones y pláticas con la Secretaría de Transporte del Estado con la finalidad de colaborar y establecer las reglas de operación entre las dos instituciones, determinar el tipo de tecnologías que debían tener las credenciales para que estas fueran compatibles con el Sistema de transporte, cumpliendo así con unos de los principales objetivos del proyecto, el de consolidar múltiples credenciales en una sola.

Posteriormente se inició con el diseño del nuevo Sistema de Credencialización y la inclusión de los nuevos tipos de trámite en el sistema de Ventanilla Única de Servicios, una vez terminado comienza el proceso de desarrollo, pruebas y liberación respectivamente.

Todo esto se llevó a cabo bajo la metodología SCRUM y el uso de micro servicios que permiten vincular los sistemas de información a través de servicios web.

Parte de las funcionalidades en los sistemas mencionados son las siguientes:

Ventanilla Única de Servicios

- Inclusión de dos nuevos tipos de trámite “Credencial nueva para primer ingreso” y “Reposición de credencial” para realizar la solicitud, así como un listado de validaciones predefinidas.
- Carga de aranceles institucionales por tipo de trámite.
- Validación de pagos recibidos.
- Trazabilidad e histórico de los trámites

Sistema de Credencialización Universitaria

- Autenticación bajo esquema de seguridad
- Listado de trámites pendientes de atender
- Pre visualización de las credenciales para ser autorizadas o rechazadas
- Mensajes de confirmación y resumen de los trámites autorizados y

rechazados

- Aplicativo de escritorio para la impresión, relación de listado para vinculación, impresión por excepción.
- Soporte para impresión a dos modelos diferentes de impresora.
- Módulo para la cancelación de credenciales

Vinculación con el Sistema de Transporte del Estado

- Envío de información necesaria para la activación de las credenciales, y puedan ser utilizadas como monedero electrónico.
- Envío de listado de credenciales canceladas para desactivarlas



Fig 1. Diseño de nueva credencial universitaria



Fig 2. Credencial con tecnologías de proximidad, contacto y banda magnética

Dicho proceso se muestra a continuación gráficamente:



Fig 3. Acceso al sistema mediante credenciales propias



Fig 4. Sistemas en SIIAU



Fig 5. Tipos de trámite en Ventanilla Única de Servicios

Requisitos-Ayuda	Número de trámite	Trámite	Monto	Adeudo	No. copias	Estatus	Centro Universitario	Carrera	Fecha de solicitud	
	9446	Credencial Nueva para Primer Ingreso	\$73.00	\$73.00	1	SOLICITADO	OFICINA RECTORIA SUV	LTEI	16/01/2019 13:25:19	
Adeudo Total:			\$0							
(1 de 1) 1 10										
Requisitos-Ayuda	Número de trámite	Trámite	Monto	Adeudo	No. copias	Estatus	Centro Universitario	Carrera	Fecha de solicitud	
	9446	Credencial Nueva para Primer Ingreso	\$73.00	\$73.00	1	PENDIENTE DE PAGO	OFICINA RECTORIA SUV	LTEI	16/01/2019 13:25:19	

Fig 6. Carga de aranceles institucionales

	Número de trámite	Trámite	Código del alumno	Nombre del alumno	Correo del alumno	No. copias	Carrera	Estatus	Fecha cambio de Estatus	Código	Observaciones
<input checked="" type="checkbox"/>	9446	Credencial Nueva para Primer Ingreso	304202163	BRIONES ISLAS ANA NAYELI	none@msn.com	1	LTEI	PENDIENTE DE PAGO	16/01/2019 13:40:12		
<input checked="" type="checkbox"/>	9448	Credencial Nueva para Primer Ingreso	304379888	NAPOLÉS LOPEZ JULIO CESAR	none@msn.com	1	LTEI	PENDIENTE DE PAGO	16/01/2019 14:28:19		
<input checked="" type="checkbox"/>	9447	Credencial Nueva para Primer Ingreso	211722555	GONZALEZ GONZALEZ MARIA DEL ROSARIO	none@msn.com	1	TURI	PENDIENTE DE PAGO	16/01/2019 14:20:26		

Fig 7. Validación de pagos en Ventanilla Única de Servicios

A continuación, se muestra la evidencia del uso de micro servicios para la interoperabilidad entre los dos sistemas internos de la Institución:

La Universidad de Guadalajara cuenta con una cantidad extensa de sistemas diseñados, desarrollados e implementados sobre una plataforma de desarrollo propia de la CGTI, la cual permite cumplir diversas necesidades de la institución, debido a que se tienen sistemas y/o aplicaciones hechos a la medida.

La arquitectura en la cual se encuentran desarrolladas las aplicaciones de la Universidad de Guadalajara son desarrollos propios acoplados mediante un esquema de seguridad semejante a un Single sign on, el cual permite que varios aplicativos se encuentren activos mediante una misma credencial.

Continuación del proceso en el Sistema de Credencialización Universitaria:



Fig 8. Módulo Listado de trámites para el seguimiento.

Se incorporó una funcionalidad para pre visualizar los datos que van impresos en la credencial, antes de enviar un lote de impresión, con la finalidad de validar que cuenta con todos los datos necesarios y la fotografía cuente con el requerimiento establecido por la Coordinación General de Control Escolar -CGCE-, para que la fotografía sea válida debe ser tomada con fondo blanco en una superficie plana, el rostro de la persona debe estar enfocada de la nariz hacia las orejas, debe ser frontal a la persona, la cual no genere ningún tipo de sombra o mancha en la fotografía.

Además, el sistema muestra un mensaje con el resumen de trámites pre visualizados y cuantos de ello fueron autorizados y rechazados antes de ser enviados a impresión.



Fig 9. Pre visualización y resumen de trámites

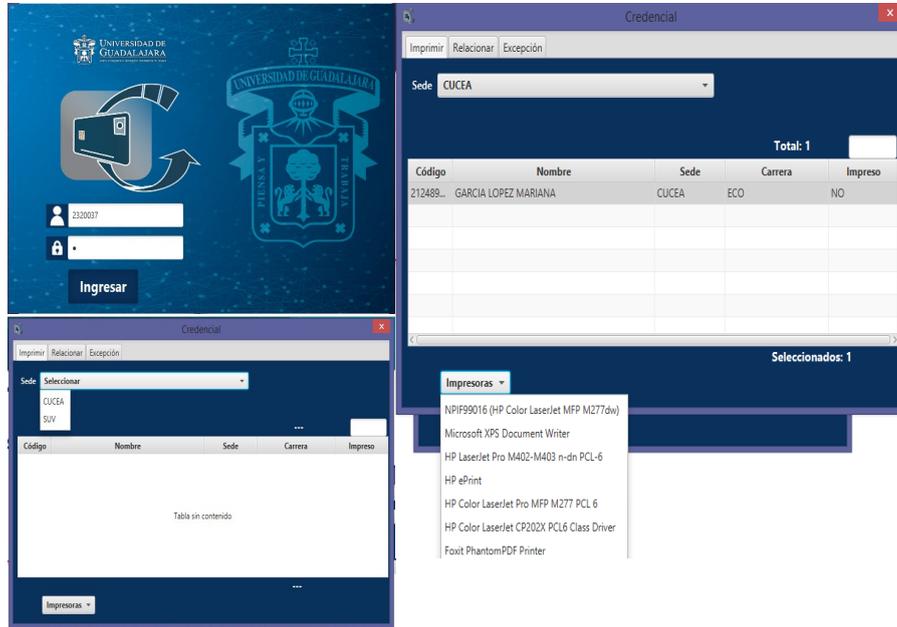


Fig. 10 Aplicativo de escritorio para la impresión

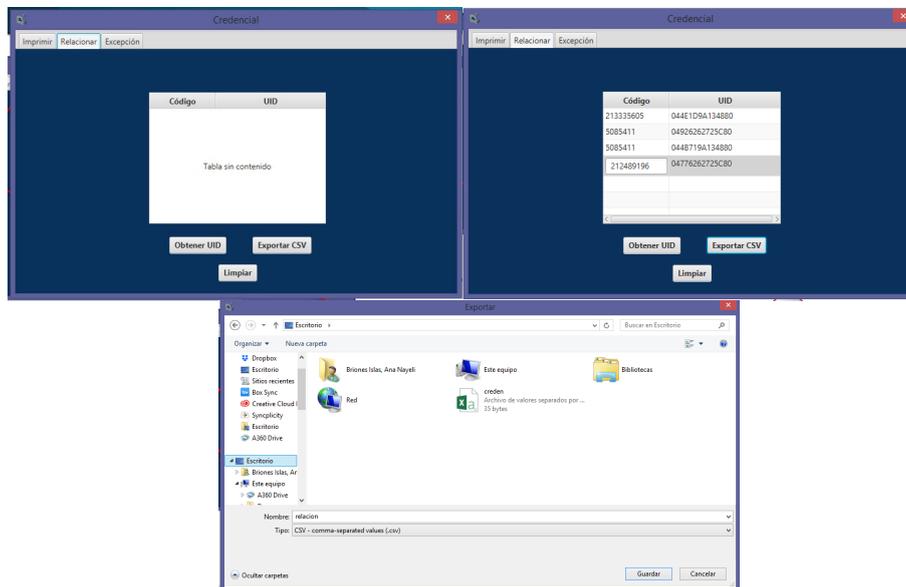


Fig 11. Aplicativo de escritorio para la relación código-UniqueID

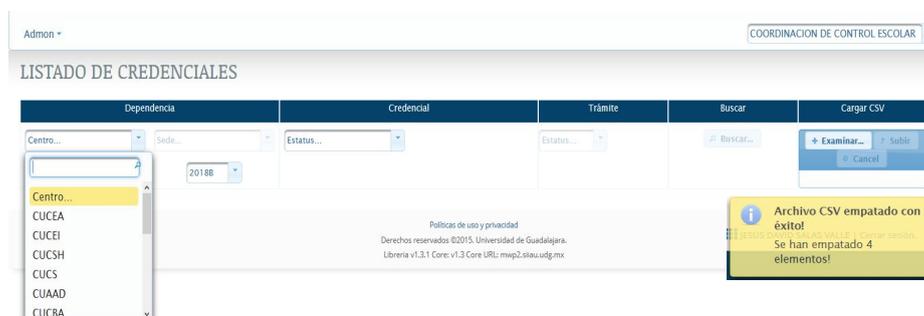


Fig 12. Caso de éxito- Vinculación de código con UniqueID

4. Aspectos críticos y relevantes a resaltar y detallar

Algunos de los aspectos relevantes que podemos mencionar son los siguientes:

El primer aspecto tiene que ver con el trabajo que se realizó de manera colaborativa entre las diferentes áreas administrativas y operativas de la Institución, para ello se requirió de la experiencia y propuestas innovadoras del personal técnico de la CGTI, CGCE, Dirección de Finanzas, Federación de Estudiantes Universitarios.

Otro aspecto importante fue el uso de tecnologías y herramientas actuales para lograr la comunicación de manera transparente y eficiente entre los sistemas internos de la Universidad, a través del uso de micro servicios que fortalecen la interoperabilidad, y nos permitió realizar el empate de código de alumno con los identificadores únicos de las credenciales, con lo cual se tiene la certeza de que un alumno solo cuenta con una credencial activa y vigente, para su validación los sistemas realizan consultas directas a la base de datos del Sistema Integral de Información y Administración Universitaria. Aunado a ello la decisión de adquirir cierto tipo de impresoras nos garantiza que la posibilidad de falsificación es muy baja ya que la tecnología inmersa en las credenciales es Mifare Desfire EV1 con capacidad de 4K, de la mano con un diseño propio y original que va pre impreso en las credenciales tipo CR-80 que se adquieren.

El tercer aspecto relevante fue la colaboración tan estrecha que se mantuvo con la Secretaría de Transporte, esto con la finalidad de mantener armonía entre la tecnología que utiliza todo el sistema de transporte del estado y las credenciales de la Universidad, de esta manera un alumno activo y vigente de la Institución puede acudir a los módulos del estado y realizar la recarga de saldo o bien algún tipo de programa social para el beneficio de traslado a mitad de precio o gratis, con ello el alumno que es usuario del transporte eléctrico público no necesita adquirir una tarjeta diferente a la credencial universitaria.

5. Resultados obtenidos y su impacto

Con la implementación de este proyecto se ha beneficiado a cerca de 288,000 alumnos

de la Institución, al otorgarles una credencial con tecnología les permite consolidar varias credenciales en una sola y acceder a servicios dentro de sus campus, así como al personal administrativo que ejecuta el proceso de credencialización reduciendo el tiempo de atención y los costos que dicho proceso conlleva.

Hasta el momento la numeralia de los trámites realizado y atendidos es la siguiente:

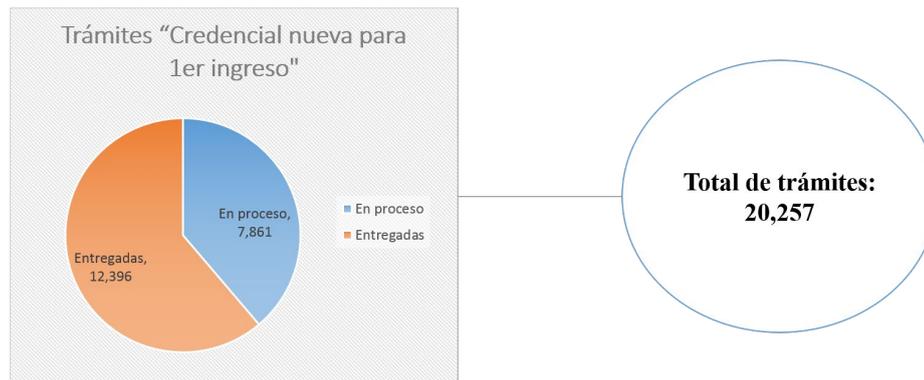


Fig. 13. Trámites de tipo "Credencial nueva para primer ingreso"



Fig 14. Trámite de tipo "Reposición de credencial"

6. Aprendizajes

Tabla 1. Aprendizajes

Universidad de Guadalajara		
Coordinación General de Tecnologías de Información		
Aprendizajes		
	¿Qué Trabajo Bien?	¿Qué Puede Ser Mejorado?
Definición de Requerimientos y Administración	La estrategia para el levantamiento de requerimientos de manera cercana a los interesados y afectados de los sistemas, se identificó al personal con el conocimiento técnico y funcional, impulsando de manera considerable el diseño del proyecto.	Dimensionar de manera correcta el tiempo y esfuerzo que implica cualquier cambio en los requerimientos ya en el proceso de implementación o desarrollo del proyecto, como un cambio en la tecnología seleccionada.
Definición de Alcance y Administración	Se cumplió con los requerimientos establecidos por el cliente, incluso se atendieron algunos otros que se sumaron en el proceso de desarrollo, que surgieron a partir de la implementación de las funcionalidades de impresión.	Definir los paquetes de liberación de las fases consideradas, así como la priorización de las mismas.
Disponibilidad de Recursos Humanos, Equipos de Desarrollo y Desempeño	La disposición de las personas asignadas como equipo de trabajo, fue significativa para el avance del proyecto y el cumplimiento de los entregables en tiempo y forma.	La cantidad de personas asignadas fue poca para el tiempo que se estableció como fecha fin del proyecto.
Administración de Comunicación	Se trabajó bajo un plan de comunicación interno y externo a la CGTI, esto ayudo a delimitar y aclarar los flujos y roles establecidos.	En ocasiones era tanta la información que debía enviarse, que los flujos de comunicaciones no siempre fueron tan efectivos como se esperaba.
Administración de Grupos de Interés	Se mantuvo informado a todo interesado en el proyecto, con su respectivo enfoque y grado de interés, lo cual nos permitió contar con la toma de decisiones de manera rápida y eficiente.	
Informes ejecutivos	Se realizó el informe ejecutivo cada 15 días ante el Comité de Proyectos, cada semana se tenían reuniones con el Cliente para revisar pendientes y mostrar avances, y cada dos días el equipo interno se reunía para establecer como solventar los problemas que se	El formato en que se presentaban los informes puede ser más dinámico y visual.

	presentaban.	
--	--------------	--

Referencias

[1] Paola Navarrete. (2013). Numeralia Institucional. 17 de mayo 2019, de Universidad de Guadalajara Sitio web: <http://copladi.udg.mx/estadistica/numeralia>

[2] Iván Serrano Jáuregui. (2019). Presentan credencial inteligente que permitirá a estudiantes de UdeG utilizar diversos sistemas de transporte. 17 de mayo 2019, de Universidad de Guadalajara Sitio web: <http://udg.mx/es/noticia/presentan-credencial-inteligente-permitira-estudiantes-udeg-utilizar-diversos-sistemas>

A teoria e a prática da gestão dos documentos digitais: o caso da Universidade Federal de Santa Maria

Marcelo Lopes Kroth^a, Gustavo Zanini Kantorski^a, Marcos Vinicius Bittencourt de Souza^a, Daiane Regina Segabinazzi Pradebon^a, Débora Flores^a, Neiva Pavezi^a

^a Universidade Federal de Santa Maria,
97.105-900, Campus Universitário, Camobi, Santa Maria, RS, Brasil
marcelo.tuco@ufsm.br, gustavo@ufsm.br, marcos.cpd@ufsm.br,
daianepradebon@gmail.com, deboraufsm@gmail.com, neivapavezi@gmail.com

Resumo. As universidades da América Latina estão passando por um processo rápido de transformação digital e de forma irreversível. Cada vez mais, estamos produzindo um grande volume de documentos digitais nas nossas instituições sem a devida atenção à gestão arquivística. A atual inobservância de requisitos para a produção de “bons” documentos digitais pode nos levar a um futuro onde a autenticidade do que está sendo produzido atualmente seja questionada. Outro fator que preocupa os pesquisadores da área é a perda de contexto dos documentos ou, até mesmo, a perda total do documento digital. Os desafios para a presunção de autenticidade e preservação em longo prazo dos documentos digitais são muitos, porém existem normas internacionais de gestão arquivística e preservação digital que norteiam os esforços que devem ser feitos para que se tenha um mínimo de segurança para a produção de documentos digitais em nossas instituições. Este trabalho descreve as ações realizadas na Universidade Federal de Santa Maria para a gestão, preservação e acesso em longo prazo aos documentos arquivísticos digitais. O objetivo é a construção um novo padrão multidisciplinar de transparência e modernização dos atos administrativos com responsabilidade, sustentabilidade e inovação. É apresentado um modelo contendo componentes importantes para presunção de autenticidade e garantia de acessibilidade a gerações futuras. A experiência compartilhada pode ser aproveitada ou servir de inspiração a outras universidades.

Palavras Chave: Gestão de documentos. Documentos digitais. Preservação digital.

Eixo temático: *Respondiendo a la universidad cambiante de la era digital mediante desarrollos rápidos y flexibles.*

1 Introdução

Os documentos produzidos por instituições públicas, no exercício de suas funções e atividades, são instrumentos fundamentais para a tomada de decisão, comprovação de direitos individuais e coletivos e para o registro da memória social.

Os documentos digitais estão presentes, cada vez mais, nos processos de trabalho. Grande parte da produção documental é feita pelos sistemas de negócio informatizados com insuficiente, ou totalmente ausente, atenção à gestão arquivística. As universidades, principalmente da América Latina, ainda não estão preparadas para superar os desafios para a uma efetiva gestão dos documentos digitais. A transformação digital pela qual estamos passando, demanda novas competências nas nossas instituições. Na maioria das vezes, os documentos digitais são gerenciados e preservados por profissionais da área de TIC, fato que pode levar à perda de contexto e autenticidade ou, até mesmo, à perda total do documento digital [1].

Segundo Aída Luz Mendoza Navarro [2], em entrevista à revista *Documentación*, os principais desafios relativos aos documentos digitais são a preservação e a demonstração em longo prazo da autenticidade.

La política pública, los archivos y la legislación constituyen tres elementos indisolubles para garantizar los derechos y obligaciones del Estado en su relación con los ciudadanos. La transparencia y el acceso a la información, así como todas las nuevas tendencias a través de las cuales los estados ponen a disposición la información a través de las TIC, solo serán exitosas si los tres elementos son asumidos por quienes tiene poder de decisión con la debida firmeza y convencimiento.

A característica inerente de construção do documento digital apresenta uma complexidade adicional para presunção de autenticidade e preservação em longo prazo. A dificuldade em relação à presunção de autenticidade se deve em razão de que os documentos digitais podem ser facilmente duplicados, distribuídos, renomeados, reformatados ou convertidos, além de poderem ser alterados e falsificados com facilidade, sem deixar rastros aparentes [3]. As principais questões que envolvem a preservação digital em longo prazo são a falta de normalização e obsolescência dos formatos de arquivo.

A digitalização dos documentos tornou-se a “saída mágica” para grande parte das instituições para dar agilidade aos processos de negócio. Porém, caso fosse feita a gestão dos documentos utilizando tabelas de temporalidade e eliminando o que é permitido, os gastos com a digitalização poderiam ser minimizados, bem como o acúmulo desnecessário de documentos.

Para que seja possível assegurar a autenticidade e preservação em longo prazo dos documentos nato digitais é fundamental observar os requisitos para a produção de “bons” documentos, ou seja, os processos de negócio informatizados devem utilizar como suporte um sistema de gerenciamento de documentos, além de um repositório digital.

A implantação de política e de procedimentos de gestão arquivísticos é imprescindível no mundo digital e, para tal, já existem requisitos, padrões e recomendações internacionais.

A maioria dos países da América Latina aprovou normas sobre a aplicação da tecnologia da informação nos arquivos (Tabela 1), mas a perda de informações não

pode ser evitada com uma lei, é preciso trabalhar com outros elementos, sendo que um deles são requisitos arquivísticos.

Tabela1. Legislação dos países da América Latina

País	Legislação
Argentina	Ley del Derecho de Acceso a la Información Pública - Ley 27.275, de 14/9/2016
Brasil	Lei de Acesso à Informação - Lei nº 12.527, 18/11/2011
Chile	Ley de Transparencia de la Función Pública - Ley nº 20.285, 20/8/2008
Colombia	Ley de Transparencia y de Acceso a la Información - Ley nº 1.712, 6/3/2014
Costa Rica	Ley del Derecho de Petición de Información - Ley nº 9.097, 26/10/2012
Guatemala	Ley de Acceso a la Información Pública - Decreto nº 57, 23/9/2008
Honduras	Ley de Transparencia y Acceso a la Información Pública - Decreto nº 170, 27/11/2006
México	Ley General de Transparencia y Acceso a la Información Pública - Decreto de 16/4/2015
Nicaragua	Ley de Acceso a la Información Pública - Ley nº 621, 16/5/2007
Panamá	Ley de Autoridad Nacional de Transparencia y Acceso a la Información - Ley nº 33, 25/4/2013
Paraguay	Libre Acceso Ciudadano a la Información Pública y Transparencia Gubernamental - Ley nº 5.282, 18/9/2014
Perú	Ley de Transparencia y Acceso a la Información Pública - Ley nº 27.806, 13/7/2002
República Dominicana	Ley General de Libre Acceso a la Información Pública - Ley nº 200, 28/7/2004
Uruguay	Ley de Acceso a la Información Pública - Ley nº 18.381, 7/11/2008

O Projeto InterPARES, liderado por Luciana Duranti, foi a principal fonte de desenvolvimento da base de conhecimento teórico-metodológico essencial para a preservação de longo prazo de documentos arquivísticos digitais autênticos. Um dos resultados foi concluir, de forma empírica, que é impossível preservar os documentos digitais devido à sua construção inata, só sendo possível preservar a capacidade de reproduzi-los [4]. A reprodução de documentos digitais tornou-se o único meio que os usuários humanos podem confiar para acessá-los novamente após a primeira vez que são salvos, independentemente de quanto tempo eles vão existir. Assim, a preservação da capacidade de reproduzir, tornou-se a pedra angular da preservação digital [5]. Neste cenário, a presunção de autenticidade dos documentos arquivísticos digitais deve se apoiar na evidência de que eles foram mantidos com uso de tecnologias e procedimentos que garantam a sua identidade e integridade; ou que, pelo menos, minimizem os riscos de modificações dos documentos a partir do momento em que

foram salvos pela primeira vez e em todos os acessos subsequentes [3].

Além disso, essa presunção baseia-se na confirmação da existência de uma cadeia de custódia ininterrupta²⁵, desde o momento da produção até a transferência para a instituição arquivística responsável pela sua preservação. Caso essa cadeia de custódia seja interrompida, o tempo em que os documentos não estiveram sob a proteção do seu produtor ou sucessor pode causar dúvidas sobre a sua autenticidade [3].

Para garantir a integridade dos documentos digitais, é fundamental a observância dos conceitos de forma fixa, conteúdo estável, forma documental armazenada ou manifestada, assim como a fixidez da informação em seu suporte de forma indissociável. A utilização da forma documental manifestada é uma das características do documento arquivístico que mais impactam na percepção humana quanto à sua confiabilidade. Talvez por isso, tem-se orientado a escolha do formato PDF/A (formato de arquivo que assegura o acesso em longo prazo de documentos eletrônicos) para a produção de documentos arquivísticos digitais, por sua baixa exigência de recursos tecnológicos para sua apresentação e sua perspectiva de perenidade [6]. Para a verificação de fixidez são utilizados algoritmos de *hash* a criação de um código a partir de um objeto digital. Se o código criado em um ponto é idêntico ao código criado pelo mesmo algoritmo em um momento posterior, isso indica que o objeto não se alterou durante esse íterim [7].

Outra grande preocupação é a preservação dos documentos em longo prazo, observando os impactos das mudanças tecnológicas, incluindo o suporte a novas mídias de armazenamento e formato de dados e ainda uma comunidade de usuários em constante transformação. Nesse contexto, os repositórios digitais vêm desempenhando um papel importante na construção de um espaço arquivístico digitais responsável pela guarda confiável de documentos digitais.

Segundo o CONARQ [8], um repositório digital de documentos arquivísticos (RDC-Arq) é um repositório digital que armazena e gerencia documentos arquivísticos digitais nas fases corrente, intermediária e permanente. Como tal, esse repositório deve ser capaz de gerenciar os documentos e seus metadados (dados estruturados que descrevem o documento) de acordo com as práticas e normas da Arquivologia, especificamente relacionadas à gestão documental, descrição arquivística multinível e preservação; e resguardar as características do documento arquivístico, em especial a autenticidade e a relação orgânica entre os documentos.

Esse artigo apresenta as ações que a Universidade Federal de Santa Maria tem feito com o objetivo de efetivar a gestão dos documentos digitais na prática. O texto está estruturado em 3 seções, sendo a Seção 1 de cunho introdutório, a Seção 2 relata os resultados alcançados até então e, por fim, a Seção 3 apresenta as considerações finais.

2 Resultados

Em 2015 foi criada uma comissão interdisciplinar com a finalidade de elaborar um projeto de informatização da gestão, preservação e acesso aos documentos arquivísticos produzidos pela universidade. Além da criação do projeto, a comissão orientou as

²⁵ Linha contínua de custodiadores de documentos arquivísticos pela qual se assegura que esses documentos são os mesmos desde o início e, portanto, são autênticos

adequações nos sistemas para atender às recomendações e normativas internacionais. O objetivo foi ampliar, de forma incremental, a produção de documentos arquivísticos nato digitais atendendo os requisitos necessários para garantir a confiabilidade, integridade e presunção de autenticidade em longo prazo.

Foi elaborado um modelo conceitual utilizado para o desenvolvimento do projeto que pode ser utilizado por outras universidades da América Latina, pois foi embasado em padrões e normas internacionais, como as orientações do Projeto InterPARES e o modelo de referência OAIS [9].

Foi utilizado como base o cenário onde o sistema de negócio desenvolvido pela UFSM - Sistema de Informações para o Ensino (SIE) - incorpora as funcionalidades de um SIGAD e interopera com um RDC-Arq [10] (Fig; 1).

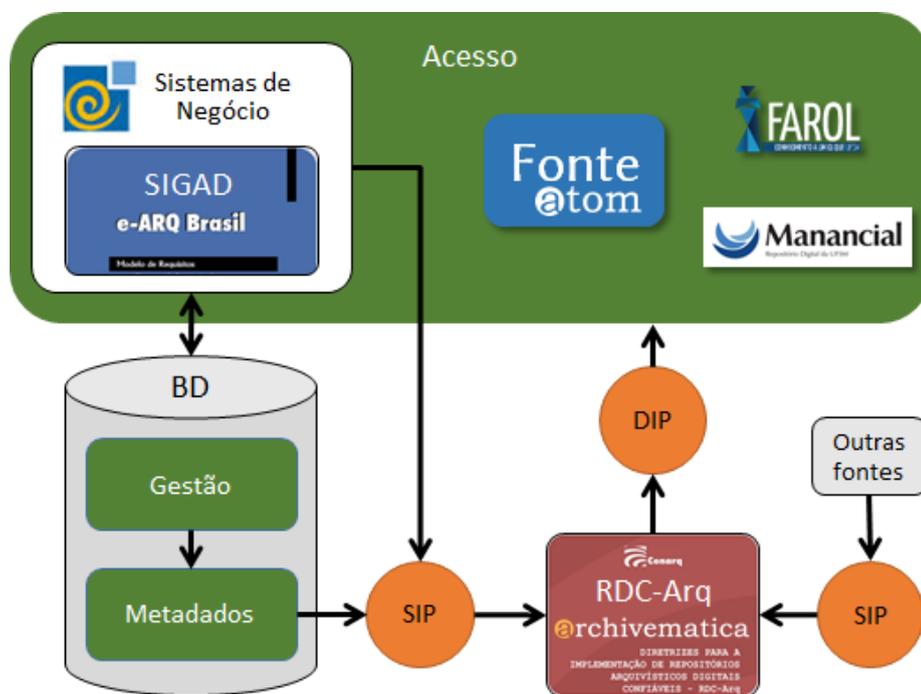


Fig. 1. Visão geral do projeto de gestão de documentos da UFSM

Por exemplo, através de funcionalidades incorporadas ao sistema de negócio, o módulo acadêmico faz a produção do documento nato digital. Os demais requisitos de gestão arquivística ficam a cargo de um núcleo comum e podem ser utilizados no gerenciamento dos demais documentos. A escolha foi natural devido à estrutura do SIE, que possui um núcleo comum utilizado por todos os sistemas. Assim, os novos requisitos desenvolvidos poderão ser absorvidos tanto pelos sistemas administrativos, quanto pelos sistemas da área acadêmica de forma gradual.

Baseado nas interfaces externas do ambiente OAIS, os Sistemas de Negócio fazem o papel de Produtor, que realiza o recolhimento dos documentos através de SIPs gerados a partir do modelo de metadados. O mecanismo proposto garante a custódia

ininterrupta dos documentos, no qual as funcionalidades nos Sistemas de Negócio enviam SIP diretamente ao repositório. Caso essa cadeia de custódia seja interrompida o tempo em que os documentos não estiveram sob a proteção do seu produtor ou sucessor pode causar dúvidas sobre a sua autenticidade [3].

O modelo é formado por diferentes consumidores, que são sistemas que darão acesso às informações preservadas no repositório. O acesso é feito através de DIP, que pode ser requisitado tanto pelo sistema de gestão, quanto pelas plataformas de acesso. Atualmente a UFSM possui, além do próprio Sistema de Gestão (SIE), as plataformas de acesso: Fonte (documentos digitais e digitalizados); Farol (vídeos e áudios focados no *streaming* de transmissão ao vivo ou gravadas); e Manancial (produção técnica e científica). O desenho projetado permite a inclusão de novas plataformas, bem como a disponibilização das informações para ambientes externos à instituição.

A integração do SIGAD com o Arquivo Digital Permanente inicia no controle da produção do documento e seus respectivos metadados, acompanha todo o seu a trâmite até a sua destinação final, além de continuar a realizar ações de preservação, visando garantir o acesso às futuras gerações. Os documentos digitais em fase permanente são dependentes de um RDC-Arq para o tratamento técnico adequado, incluindo classificação, descrição e acesso, de forma a assegurar a manutenção da autenticidade e da relação orgânica dos documentos. Optou-se pela adoção do Archivematica, que é software livre, código aberto e está em conformidade com o modelo de referências OAIS.

2.1 Plano de Classificação e Tabela de Temporalidade dos Documentos

A informatização da gestão dos documentos arquivísticos na UFSM teve como base a classificação dos documentos produzidos no Sistema de Informações para o Ensino (SIE), ação conjunta desenvolvida pelos setores de arquivo e TIC da instituição, com o assessoramento de uma equipe multidisciplinar, formada essencialmente por arquivistas e analistas de TIC.

A classificação e avaliação de documentos é um procedimento complexo, o qual exige a sistematização de atividades e a atribuição de responsabilidade de cada envolvido. O resultado desse processo culmina na adequada eliminação do documento ou na destinação para guarda permanente, visando a garantia da integridade e acesso a longo prazo dos documentos, principalmente aos digitais.

O plano de classificação e a tabela de temporalidade dos documentos passaram a ser utilizados pelo setor de arquivo da universidade no início do ano de 2017. Desde então, o setor tem autonomia para contemplar o contexto documental da universidade no sistema, através do cadastro de tipos documentais, determinando a sua classificação e a definição dos prazos de guarda (Tabela 2).

Tabela 2. Exemplo do Plano de Classificação e Tabela de Temporalidade da UFSM

Cód.	Assunto	Prazos de Guarda		Destinação Final
		Corrente	Interm.	
100	ENSINO SUPERIOR			
120	Cursos de graduação (inclusive na modalidade a distância)			
125	Vida acadêmica dos alunos dos cursos de graduação			
125.3	Avaliação acadêmica			
125.33	Registro de conteúdo programático ministrado, rendimento e frequência – diários de classe, listas de frequência, relação de notas e resumo semestral.	10 anos	10 anos	Eliminação
	Diário de classe de curso de graduação	0,5 anos	-	Guarda Permanente
	Processo de justificativa de faltas de aluno de graduação	10	10	Eliminação

Até maio de 2019, aproximadamente 200 mil documentos foram produzidos e classificados através do sistema informatizado, incluindo os Diários de classe, Portarias, Processos, Comprovantes, etc. Na Tabela 2 é possível observar o código de classificação, o tipo documental e a quantidade dos documentos mais produzidos já classificados de acordo com o plano de classificação. Estimando uma média de três páginas por documento, pode-se inferir que mais de 600 mil páginas deixaram de ser impressas e armazenadas nos arquivos da universidade, com requisitos de autenticidade e preservação.

Tabela 3. Quantidade dos principais documentos classificados na UFSM

Cód. Classe	Tipo Documental	Quant.
125.33	Diário de classe de curso de graduação	37,4 K
125.23	Processo de dispensa de disciplina de graduação	26,0 K
029.21	Portaria de autorização de afastamento eventual no país	18,6 K
	Portaria de autorização de afastamento capacitação no país	11,8 K
052.21	Comprovante de transferência de créditos entre usuários	11,5 K
134.34	Diário de classe de curso de pós-graduação stricto sensu	10,3 K

A Figura 2 mostra a tela com o plano de classificação e a tabela de temporalidade no sistema de gestão. Além das informações específicas para a gestão documental, outras

informações foram adicionadas ao plano para auxiliar na produção e padronização dos documentos digitais, tais como, restrição de acesso, workflow, mapeamento como a classificação antiga, texto padrão, etc. Dessa forma, também foi possível atender grande parte dos requisitos do Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos (e-ARQ Brasil) [11].

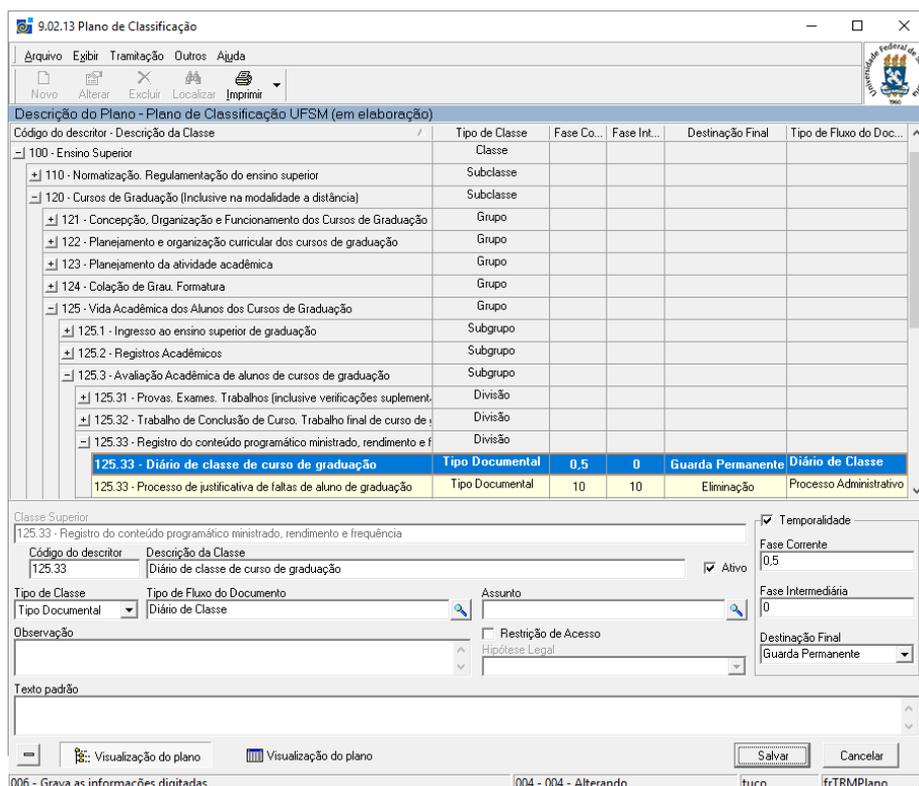


Fig. 4. Tela do sistema com o Plano de Classificação e a Tabela de Temporalidade

É importante ressaltar que a questão que pesa sobre a eliminação de um documento digital não se refere a ter ou não ter espaço suficiente para armazenamento, mas sim sobre se sua preservação (e o custo decorrente de mantê-lo) é necessária [6].

Documento

Diário de classe da disciplina MATEMÁTICA INSTRUMENTAL (UDSM1064) turma 10
Curso Superior de Tecnologia em Gestão Ambiental (3000) ano 2017 1. Semestre

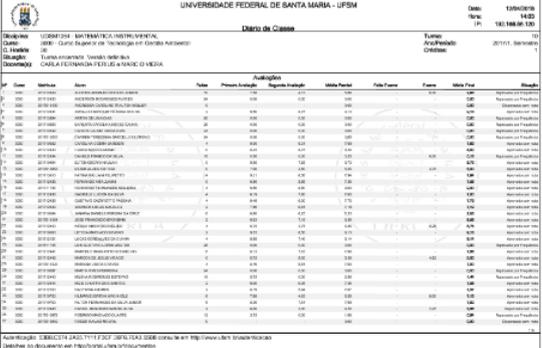
Descrição
 Diário de classe da disciplina MATEMÁTICA INSTRUMENTAL (UDSM1064) turma 10 curso Curso Superior de Tecnologia em Gestão Ambiental (3000) ano 2017 período 1. Semestre. É o documento oficial utilizado para o controle da frequência e aproveitamento dos alunos, além do registro do conteúdo ministrado nas aulas. É gerado no Portal do Professor, sob a responsabilidade do(s) professor(es) de cada turma.

Procedência
 DEPARTAMENTO DE ENSINO - POLITÉCNICO - 26.04.00.00.0.0 (Unidade administrativa)

Classificação
 125.33 - Registro do conteúdo programático ministrado, rendimento e frequência

Tipo documental
 Diário de classe de curso de graduação

Primeira página do documento



Data de produção
12/04/2018

[diario_classe_UDSM1064_T_10_A_2017_P_1_Semestre.pdf \(354 KB\)](#)

Fig. 3. Exemplo de documento nato digital classificado no sistema de gestão

A Fig. 3 mostra um exemplo de documento nato digital (Diário de Classe) produzido pelo sistema de gestão já classificado conforme o plano de classificação e tabela de temporalidade.

2.2 Modelo de Metadados

O modelo proposto, os metadados são capturados a partir das informações geradas por eventos no sistema de gestão. Os metadados ficam separados logicamente dos dados de gestão e podem ser utilizados de forma independente.

- *Metadados* abrange o registro do metadados em si, incluindo sua categoria e um tipo de especificação. O modelo suporta mais de uma especificação. Além do e-ARQ Brasil, o modelo também suporta a especificação ISAD(G), ePMG e Dublin Core, além de uma especificação própria da instituição.

- *Metadados Relações* permite que metadados de especificações distintas sejam relacionados. O relacionamento é sempre realizado aos pares.
- *Metadados Agentes* identifica o agente que realizou determinado evento, pelo qual é responsável, em um documento arquivístico. O nome do agente, sua autorização de acesso ao documento, credencias de autenticação (senha), relação e ou papel do agente e situação do agente são registradas.
- *Metadados Eventos* registra o evento realizado pelo agente no documento em determinado instante de tempo. Eventos de gestão que podem ser registrados incluem a captura, tramitação, recolhimento, juntada, entre outros.
- *Metadados Instâncias* armazenam os metadados do documento conforme definido em Metadados. Os metadados são registrados em pares (chave, valor). Exemplos de metadados são: (Idioma, 'pt-BR'), (Tipo, 'Diário de Classe'), (Prazo de Guarda, '2017-06-05T14:50:34-03:00').
- *Componentes Digitais* englobam os arquivos armazenados juntamente com os metadados sobre os componentes. Além do conteúdo do arquivo, são registradas informações como a ordem do componente digital no documento arquivístico digital, o nome do arquivo, o *MIME Type* do arquivo, seu tamanho e sua fixidez. A fixidez tem como objetivo validar se o componente digital foi alterado de forma não documentada ou não autorizada.

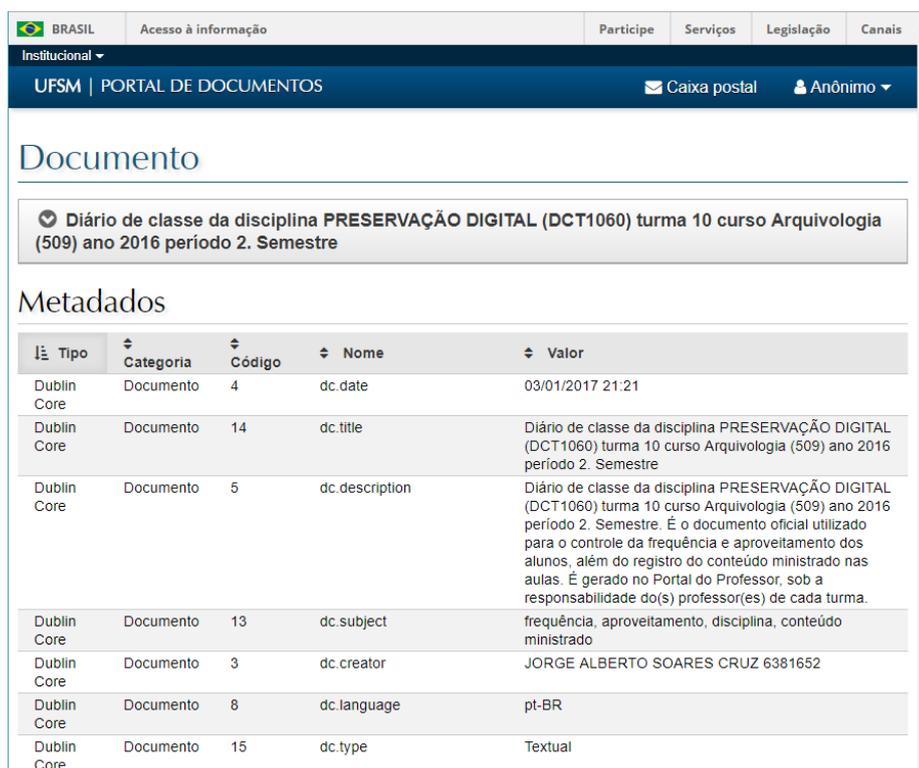


Fig. 4. Modelo de metadados da UFSM

A primeira aplicação do modelo de metadados (Fig. 4) foi no Diário de Classe da UFSM, documento oficial que possui o controle de frequência, aproveitamento dos alunos e registro de conteúdo ministrado nas aulas pelos docentes. Os documentos são de acesso público (ostensivos) e estão disponíveis para consulta no sítio institucional da universidade (<https://portal.ufsm.br/documentos>).

A quantidade de elementos de metadados é variável para cada documento. Nos primeiros semestres de utilização, o número mínimo de foi de 50 e o máximo foi 169. No total a média ficou em 65 metadados gerados por Diário de Classe.

A produção iniciou no segundo semestre de 2016 a partir da publicação da Instrução Normativa N. 03/2016 da Pró-Reitoria de Graduação, que estabelece orientações para o Diário de Classe Nato Digital no âmbito da instituição. A norma reconhece o documento digital como válido para todos os efeitos administrativos e legais, desobrigando a impressão e arquivamento nos Departamentos Didáticos. Ainda, estabelece que os arquivos gerados devam ser mantidos no Repositório Arquivístico Digital Confiável da UFSM, sendo seu prazo de guarda permanente, conforme definição da Comissão Permanente de Avaliação de Documentos (Fig.5).



The screenshot shows the 'Portal de Documentos' interface. At the top, there is a navigation bar with 'BRASIL', 'Acesso à informação', and 'Participe', 'Serviços', 'Legislação', 'Canais'. Below this is a header for 'UFSM | PORTAL DE DOCUMENTOS' with a 'Caixa postal' icon and an 'Anônimo' dropdown. The main content area is titled 'Documento' and displays the document title: 'Diário de classe da disciplina PRESERVAÇÃO DIGITAL (DCT1060) turma 10 curso Arquivologia (509) ano 2016 período 2. Semestre'. Below the title is a 'Metadados' section with a table listing various Dublin Core metadata fields.

Ícone	Tipo	Categoria	Código	Nome	Valor
	Dublin Core	Documento	4	dc.date	03/01/2017 21:21
	Dublin Core	Documento	14	dc.title	Diário de classe da disciplina PRESERVAÇÃO DIGITAL (DCT1060) turma 10 curso Arquivologia (509) ano 2016 período 2. Semestre
	Dublin Core	Documento	5	dc.description	Diário de classe da disciplina PRESERVAÇÃO DIGITAL (DCT1060) turma 10 curso Arquivologia (509) ano 2016 período 2. Semestre. É o documento oficial utilizado para o controle da frequência e aproveitamento dos alunos, além do registro do conteúdo ministrado nas aulas. É gerado no Portal do Professor, sob a responsabilidade do(s) professor(es) de cada turma.
	Dublin Core	Documento	13	dc.subject	frequência, aproveitamento, disciplina, conteúdo ministrado
	Dublin Core	Documento	3	dc.creator	JORGE ALBERTO SOARES CRUZ 6381652
	Dublin Core	Documento	8	dc.language	pt-BR
	Dublin Core	Documento	15	dc.type	Textual

Fig. 5. Tela de visualização dos metadados do documento *Diário de Classe*

3 Considerações Finais

A crescente utilização de sistemas informatizados nas instituições apresenta uma série de vantagens, mas também levanta questões complexas com relação à presunção de autenticidade e acessibilidade em longo prazo dos documentos nato digitais produzidos.

Neste artigo foi apresentado os avanços do projeto de gestão, a preservação e o acesso em longo prazo aos documentos arquivísticos digitais da Universidade Federal de Santa Maria - Brasil. O modelo proposto pode ser utilizado por outras universidades e adaptados a cada realidade, pois está embasado nas principais normas internacionais

de gestão arquivística e preservação digital.

Recomenda-se que os projetos iniciem com o foco na produção de "bons" documentos nato digitais. Para que isso seja possível, deve-se ter implantado um Plano de Classificação e uma Tabela de Temporalidade de Documentos. A produção documental também deve estar apoiada em um conjunto de metadados, requisito muito importante para garantir a autenticidade e preservação em longo prazo.

Um dos principais resultados do projeto foi a produção de mais de 110 mil documentos nato digitais, classificados e com o suporte de metadados. Estima-se que deixaram de ser impressas e armazenadas nos arquivos da universidade mais de 300 mil páginas em papel.

O projeto visa consolidar um modelo de gestão documental capaz de atender as exigências cada vez maiores de produção digital, sem negligenciar os requisitos para dar autenticidade e preservação em longo prazo do patrimônio documental da universidade.

Referencias

1. Innarelli, H. C.: Gestão da preservação de documentos arquivísticos digitais: proposta de um modelo conceitual. Tese (Doutorado) – Programa de Pós-Graduação em Ciência da Informação - Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo (2015)
2. Navarro, A. L. M. Política Pública, Archivos Y Legislación. Archivos, entre tradição e modernidade, vol. 1, pp. 167-177, São Paulo (2017)
3. Conarq: Diretrizes para a Presunção de Autenticidade de Documentos Arquivísticos Digitais. Rio de Janeiro: Arquivo Nacional (2012)
4. Duranti, Luciana. From digital diplomatics to digital records forensics. Archivaria vol. 68, pp. 39 a 66 (2010)
5. XIE, Sherry. Building Foundations for Digital Records Forensics: A Comparative Study of the Concept of Reproduction in Digital Records Management and Digital Forensics. The American Archivist, vol. 74, n. 2, pp. 576 a 599 (2011)
6. Santos, V. B. Documentos Arquivísticos Digitais: Um Descompasso entre a Teoria e a Prática no Brasil. Revista do Arquivo, Nº 6, São Paulo (2018)
7. PREMIS. Premis Editorial Committee. Data Dictionary section from PREMIS Data Dictionary for Preservation Metadata. Version 3.0 (2015)
8. Conarq: Diretrizes para a Implantação de Repositórios Digitais Confiáveis de Documentos Arquivísticos. Rio de Janeiro: Arquivo Nacional (2014)
9. CCSDS. Consultative Committee for Space Data Systems Secretariat. Reference Model for an Open Archival Information System (OAIS): Recommended Practice (CCSDS 650.0-M-2: Magenta Book), Washington, DC (2012)
10. Conarq. Cenários de uso de RDC-Arq em conjunto com o SIGAD. Rio de Janeiro: Arquivo Nacional (2015)
11. Conarq: e-ARQ Brasil: Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos. Rio de Janeiro: Arquivo Nacional (2011)

Arquitectura Basada en Microservicios para Aplicaciones en Alta Disponibilidad

Jhonny Ávila, Geovanny Campoverde, Víctor Saquicela

Dirección de Tecnologías de Información y Comunicación
Universidad de Cuenca
Cuenca – Ecuador

johnny.avilam@ucuenca.edu.ec, geovanny.campoverde@ucuenca.edu.ec,
victor.saquicela@ucuenca.edu.ec

Resumen. La creación de microservicios dentro del campo de desarrollo de software ha cobrado fuerza en los últimos años puesto que permite modularizar adecuadamente los sistemas de tal forma que los microservicios sean fácilmente integrables y mantenibles. Sin embargo, el hecho de tener muchos microservicios genera una complejidad al momento de manejar un ambiente de alta disponibilidad puesto que todo debe estar perfectamente orquestado para que no exista problemas con la integridad de los datos. En este trabajo se presenta un caso práctico donde se propone una solución tecnológica para el desarrollo de sistemas basados en microservicios en un ambiente escalable y de alta disponibilidad, configurado con herramientas open source que permiten controlar adecuadamente replicación de microservicios, índices centralizados, cache centralizada y monitoreo de eventos

Palabras Clave: Microservicios, monitoreo, cache, índices, balanceo, alta disponibilidad.

Eje temático: Respondiendo a la universidad cambiante de la era digital mediante desarrollos rápidos y flexibles.

1. Introducción

La Dirección de Tecnologías de Información y Comunicación de la Universidad de Cuenca ha venido desarrollando software durante varios años con la finalidad de presentar soluciones informáticas tanto al área académica como área administrativa. Las aplicaciones se han desarrollado siguiendo un modelo tradicional en donde un grupo de desarrolladores construía sistemas informáticos grandes que a la larga terminaban siendo aplicaciones muy difíciles de mantener debido a que no existía una adecuada modularización de las diferentes funcionalidades. Por tal motivo se ha visto la necesidad de migrar los sistemas informáticos antiguos a sistemas que utilicen nuevas tecnologías, siguiendo un marco de trabajo ágil, que sean dinámicos ante el cambio y fácilmente mantenibles.

Actualmente, las nuevas tendencias de desarrollo de software se orientan a obtener aplicaciones que presenten una alta escalabilidad, alta disponibilidad, rapidez ante las solicitudes de nuevos requerimientos y sobre todo que su mantenimiento no signifique una tarea compleja. Además, se busca que el software se encuentre modularizado adecuadamente de tal forma que los nuevos miembros de un equipo de desarrollo puedan integrarse fácilmente en el proceso de desarrollo. La mayoría de lenguajes de programación del lado del servidor como C++, Python, Java, proveen abstracciones para el desarrollo de aplicaciones en módulos, sin embargo, estos módulos comparten recursos tanto lógicos (Bases de datos, archivos, memoria) como físicos (hardware) por lo que no se pueden ejecutar independientemente uno del otro, sino se ejecutan como un solo artefacto, a estos sistemas se los denomina monolíticos. [3]

Como se indica en [1] la arquitectura de microservicios provee software simple, conformado por pequeños servicios independientes en donde cada uno ejecuta tareas específicas y se interconectan entre ellos mediante canales de comunicación. Esto, en conjunto con un ambiente de balanceo adecuado provee a los sistemas informáticos un entorno de alta disponibilidad, garantizando que los servicios siempre se encontrarán disponibles para el usuario final.

El presente documento contiene la descripción de un caso de aplicación real en la Universidad de Cuenca, en donde se ha implementado la arquitectura basada en microservicios dentro de un ambiente de alta disponibilidad.

2. Microservicios dentro de una arquitectura de alta disponibilidad

Problemática

Las aplicaciones monolíticas son componentes de software vistas como un solo bloque de código que se ejecuta en el servidor. Esto, no es deseable puesto que un pequeño cambio en el mismo podría significar una gran cantidad de tareas. Bakshi explica en [2] que las aplicaciones monolíticas grandes son mantenidas por varios equipos de desarrollo y que se necesita una minuciosa orquestación entre todos los equipos involucrados cuando se realiza un cambio o se desarrolla una nueva funcionalidad

únicamente pudiendo escalar de manera vertical.[15]

La arquitectura basada en microservicios propone una solución a la problemática general que presentan las aplicaciones monolíticas. Crear software utilizando esta arquitectura significa que se tendrán sistemas dinámicos, en donde el equipo de desarrollo pueda recibir retroalimentación de manera ágil y su respuesta ante un cambio sea rápida. Además, al estar modularizadas las funcionalidades del sistema en servicios pequeños, los diferentes equipos pueden realizar cambios sin afectar el código del resto de equipos de trabajo permitiendo escalamiento horizontal.

Una arquitectura basada en microservicios debe estar acompañada de un entorno adecuado de balanceo de carga, que garantice la disponibilidad permanente de los servicios. El poseer un ambiente de balanceo permite que se pueda replicar fácilmente aquellos microservicios que tienen mayor carga y transaccionalidad, evitando la sobrecarga del servidor y aumentando los tiempos de respuesta ante las peticiones enviadas desde los servicios externos. Es importante recalcar que el ambiente de balanceo además de garantizar la disponibilidad de los servicios debe cumplir con las reglas de integridad y consistencia de datos.

Microservicios

Como se explica en [5], un microservicio es una pequeña aplicación o componente, fácil de desarrollar y altamente escalable que puede ser fácilmente integrada con otros componentes para formar un sistema. Los microservicios utilizan canales de comunicación como API's o servicios REST para comunicarse con otros recursos como bases de datos, repositorios documentales e inclusive otros microservicios.

Entre las principales características de un microservicio se tiene que su uso puede proveer un gran modularidad a los sistemas y romper el esquema de aplicaciones monolíticas. Además, permite que los desarrolladores trabajen en diferentes funcionalidades a la vez sin afectar el trabajo de los otros equipos, aplicando frameworks de trabajo como DevOps [9]. Otra característica importante es que, al ser aplicativos modulares, se adaptan perfectamente a la arquitectura de contenedores [7], en donde cada servicio se encontrará dentro de un contenedor independiente.

Retos de balanceo de microservicios en ambientes de alta disponibilidad

Los sistemas monolíticos tienen una ventaja importante frente a la arquitectura basada en microservicios, esta es su facilidad de despliegue [4]. Por lo general, en estos sistemas es necesario desplegar un solo archivo o conjunto de archivos en un servidor de aplicaciones y se tendrá el sistema completo desplegado. De igual forma, el camino para la escalabilidad es más simple puesto que solo se necesita ejecutar múltiples copias de la aplicación completa en un ambiente de balanceo de carga, sin embargo, este escalamiento es en una sola dimensión, es decir solo se incrementa el número de copias de la aplicación, sin tomar en cuenta los componentes que tienen una mayor o menor carga de trabajo o el consumo de recursos de cada uno. Esto puede significar un incremento de coste significativo tanto en recursos físicos como lógicos y humanos para mantener el sistema estable al momento de escalar.

En la arquitectura de microservicios, al ser cada componente independiente del otro, se puede escalar individualmente destinando únicamente los recursos necesarios de acuerdo a sus necesidades, de esta forma se evita el desperdicio de recursos y se puede mantener infraestructuras más pequeñas y manejables. Sin embargo, como ya se había dicho, esta arquitectura presenta cierta complejidad adicional al momento de crear sistemas distribuidos.

Probablemente uno de los retos más importantes en esta arquitectura es encontrar la manera de implementar un mecanismo de comunicación efectivo entre los diferentes servicios garantizando así la transaccionalidad en las operaciones y la consistencia de los datos de cada componente. Para esto, es importante decidir la manera correcta de particionar al sistema en los diferentes componentes o microservicios para aprovechar todos los beneficios de esta arquitectura, sin embargo, esto también incrementa la complejidad en el desarrollo. Por lo tanto, es importante que todos los equipos implicados en el desarrollo de la aplicación tengan una fuerte coordinación y comunicación.

Arquitectura propuesta

La Dirección de Tecnologías de la Información y Comunicación (DTIC) de la Universidad de Cuenca ha visto la necesidad de mejorar sus sistemas informáticos monolíticos hacia sistemas que utilicen una arquitectura basada en microservicios, de alta disponibilidad y que sea fácilmente escalable e integrable con otros proyectos. Para solventar esto, se ha realizado la implementación de un ambiente en alta disponibilidad basada en microservicios, tal y como se observa en la Figura 1. A continuación se detalla el funcionamiento de la arquitectura propuesta.

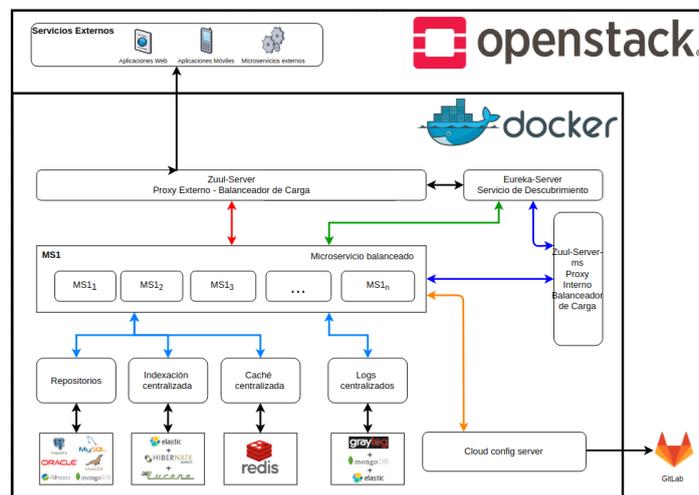


Fig. 1. Microservicio en una arquitectura de alta disponibilidad.

Cuando una aplicación web, aplicación móvil o servicio externo realiza una petición al servidor lo primero que sucede es que dicha petición se envía a un servidor proxy

(ZUUL Externo) [6], el cual indicará qué ruta debe tomar la petición. Para esto, el proxy se comunica con un servidor de descubrimiento (Eureka Server) [8], el cual contiene la información de todos los microservicios registrados, con las instancias y nombres de aquellos microservicios que han sido replicados. Aquí, es importante indicar que Eureka realiza el descubrimiento automático de los servicios cada vez que existe un nuevo microservicio o una nueva instancia del mismo.

La respuesta de Eureka permite al servidor proxy conocer qué instancias de un microservicio se tienen y en base a su algoritmo round robin [10], enviar la petición a la instancia que más crea conveniente. De esta manera, cada una de las peticiones que reciba un microservicio puede ser atendido por cualquiera de las instancias del mismo, garantizando siempre la consistencia de datos e información.

Una vez que la petición ha llegado a la instancia designada, esta puede conectarse con cualquier repositorio para obtener información tal como una base de datos, un gestor documental, etc. En caso de que la instancia desee comunicarse con otro microservicio que se encuentre alojada en la misma arquitectura de balanceo, se utilizará un servidor proxy interno (ZUUL) que a su vez se comunicará con Eureka para determinar qué ruta debe tomar hacia el nuevo microservicio.

Al mismo tiempo que el servicio se conecta con los repositorios de información se están realizando dos tareas adicionales, indexación y envío a caché. La indexación se ha configurado utilizando Elasticsearch, Hibernate Search y Lucene [11] con la finalidad de tener una indexación centralizada. De esta manera, cada vez que se ejecute una transacción desde una instancia de microservicio hacia la base de datos el índice quedará actualizado para todas las instancias del microservicio. Es importante indicar que al no disponer de indexación centralizada, las diferentes instancias de un microservicio no encontrarán estos datos en su índice hasta que se reinicien, generando la impresión de inconsistencia de datos.

Por otro lado, se ha configurado una caché utilizando un clúster de Redis [12] con 6 nodos para que funcione de manera centralizada permitiendo de esta manera que cada instancia de cada microservicio almacene su caché en un punto único.

Adicionalmente, cada instancia del microservicio estará siempre enviando logs a un servicio de logs centralizados que se ha configurado utilizando Graylog, Mongo DB y Elasticsearch [13]. La centralización permitirá que los logs enviados por cada instancia sean almacenados y organizados de tal forma que al buscar los mensajes se pueda obtener información indistintamente de la instancia de la que viene.

Finalmente, es importante indicar que cada instancia de microservicio se conectará a un servidor de configuraciones externo en donde se mantiene toda la información relacionada a comunicación con otros servicios, API's, etc.; facilitando el mantenimiento de configuraciones puesto que no se tendrá que configurar tantas veces como instancias de un microservicio existan, sino que cada instancia tomará las configuraciones desde un único punto.

Todo lo antes mencionado se encuentra en una arquitectura de contenedores basado en Docker, en donde cada instancia, cada base, cada servidor llega a ser un contenedor. Es así que gracias a esta arquitectura y al balanceo de carga se puede tener instancias de un mismo microservicio en diferentes servidores virtuales e inclusive físicos, con una adecuada comunicación y manteniendo siempre la integridad y consistencia de datos, índices, cache y logs.

Balanceo de carga

Un ambiente de balanceo de carga para microservicios debe estar adecuadamente orquestado y configurado puesto que sobre él se ejecutarán varios microservicios muchos de los cuales tendrán más de una instancia en ejecución. Para este fin, se ha implementado una arquitectura de balanceo de carga para microservicios en donde existen tres componentes principales: una herramienta de descubrimiento de microservicios, un servidor proxy interno y un servidor proxy externo.

Antes de que un servicio esté disponible en el balanceo de carga, este debe ser registrado en algún sitio de tal manera que pueda ser presentado a los servicios externos. Para este fin se ha utilizado Eureka, una herramienta de descubrimiento y registro de microservicios propuesta dentro de la arquitectura de Netflix [14]. Eureka permite que cualquier nuevo microservicio que ingrese a la red se registre automáticamente en su servidor indicando que está ahí y que puede ser consumido. Adicionalmente, con la finalidad de tener una alta disponibilidad, Eureka permite registrar varias instancias de un mismo microservicio, en donde cada una funcionará independientemente de las otras instancias sin afectar el rendimiento.

Al mismo tiempo que los microservicios están registrándose, un servidor proxy Zuul interno está actuando como canal de comunicación entre los microservicios. Este servidor interno consume la información de Eureka para conocer los microservicios que ha registrado y redireccionar peticiones entre ellos en caso de ser necesario. De esta manera, cuando un microservicio desee interactuar con otro, enviará una petición a este servidor que se encargará de enviarlo al microservicio correspondiente.

El objetivo de usar un servidor proxy Zuul es que este cuenta con un balanceador de carga llamado Ribbon [14] el cual es adecuado para manejar todas las instancias de los microservicios registrados en Eureka. Ribbon utiliza un algoritmo round robin para redireccionar las peticiones entre las diferentes instancias de un microservicio, logrando de esta manera balancear adecuadamente la carga y evitando que el microservicio se sature o deje de responder.

De la misma manera, existe un servidor proxy Zuul externo que tendrá la tarea de recibir las peticiones de los servicios externos, comunicarse con Eureka para obtener las instancias de los microservicios y enviar dicha petición a la instancia que el balanceador Ribbon haya determinado.

En la Figura 2 se encuentra la arquitectura de balanceo de carga implementada para los microservicios. Aquí se puede observar que el microservicio MS1 tiene K número de instancias, mismas que se han registrado en Eureka y que presentan esta información tanto para que el servidor proxy externo como interno puedan consumirla y direccionar adecuadamente las peticiones.

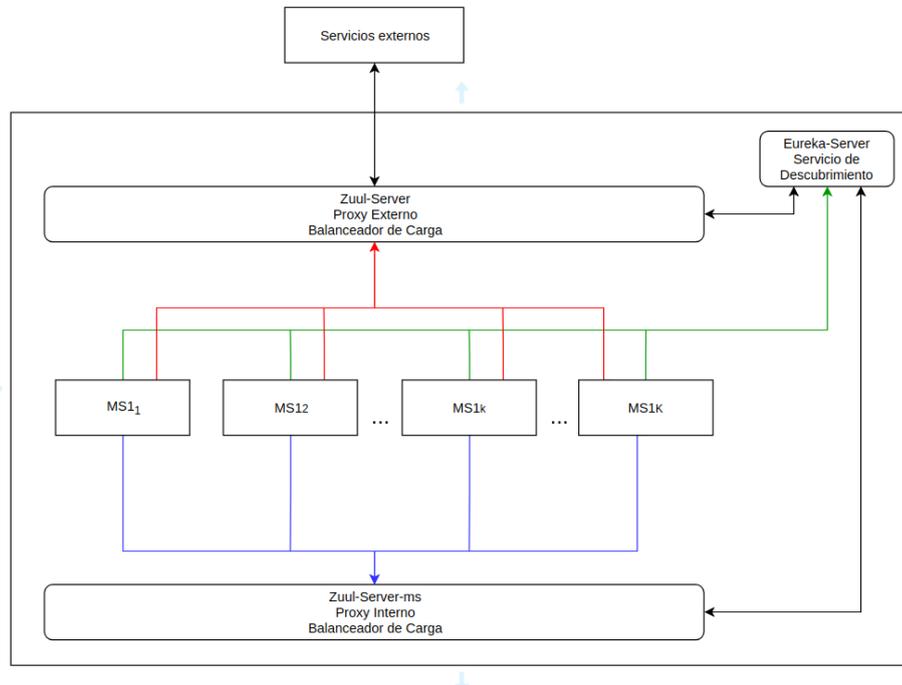


Fig. 2. Microservicio en un ambiente balanceado.

Servidor de configuraciones centralizado

Un problema común y recurrente en las aplicaciones monolíticas radica en el almacenamiento de las configuraciones de los servicios. Muchas de las veces estas configuraciones se encuentran quemadas en código o en el mejor de los casos se encuentran dentro de archivos separados que al compilarse para un ambiente de producción se empaquetarán en el mismo aplicativo haciéndolos difíciles de mantener cuando se desea realizar cambios en alguna de sus configuraciones.

En el caso de los microservicios, se generan archivos con las configuraciones que serán leídos en el arranque del contenedor. El problema radica en que al tener varias instancias de un microservicio se debe cambiar los archivos de configuración de cada instancia cuando exista un nuevo cambio, convirtiendo este trabajo en un punto de fallo pues se podría cambiar erróneamente información en alguno de los archivos generando errores o inconsistencia de datos.

Para solventar este problema la solución propuesta es utilizar un servidor de configuraciones centralizado, mismo que gracias a la arquitectura indicada anteriormente, permite que cada instancia de un mismo microservicio se conecte a un único punto de configuraciones. Dicho servidor contiene un archivo de configuración por cada microservicio, en donde se encontrarán alojadas todas las configuraciones para que el microservicio funcione adecuadamente. Así, cada vez que se genere una nueva instancia de un microservicio, esta se conectará al servidor de configuraciones y

obtendrá las mismas que el resto de sus instancias hermanas.

La ventaja principal de tener un servidor de configuraciones centralizado dentro de un ambiente balanceado es que no se da paso al fallo de evitar cambiar una configuración en un archivo de alguna de las instancias puesto que todas las instancias de un microservicio se conectarán a un único archivo y la modificación de alguna configuración en el mismo afectará a todas las instancias por igual.

En la Figura 3, se muestra el servidor de configuraciones tomando como repositorio general a un contenedor de Gitlab, en donde se mantiene un adecuado versionamiento de los archivos de configuración para los ambientes de desarrollo, pruebas y producción.

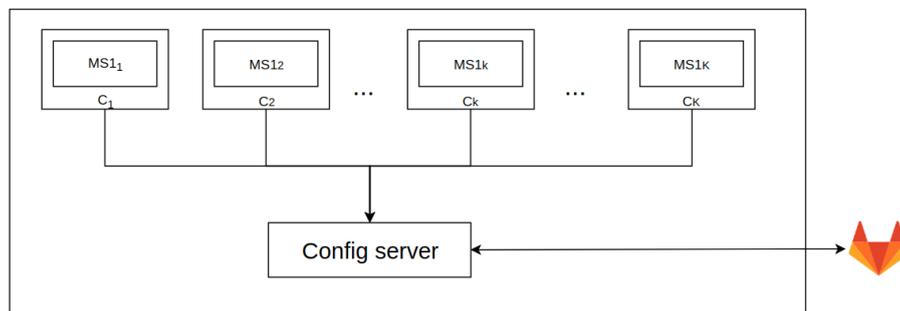


Fig. 3. Servidor de configuraciones centralizado

Como se observa que existen varias instancias del microservicio MS1, donde cada una se conecta al Config Server, el cual no es más que un pequeño servicio REST que toma el nombre del microservicio solicitante, se conecta con GitLab y busca el archivo de configuraciones correspondiente. De esta manera, cualquier cambio que se realice en los archivos de configuración alojados en Gitlab afectará a todas las instancias de los microservicios.

Indexación centralizada

Una característica importante al momento de implementar cualquier arquitectura en un sistema es que esta sea transparente al usuario final, para esto, en un ambiente de balanceo de carga, todas las réplicas de un microservicio deberían ser capaces de responder exactamente igual a una petición en un instante de tiempo si se envían los mismos parámetros.

Uno de los problemas encontrados en el balanceo de carga se presenta cuando se realizan búsquedas indexadas o de tipo full text utilizando Hibernate Search con la tecnología Lucene en los microservicios. Esto se da porque cuando se tiene más de una réplica de un microservicio que indexa los datos de la base de datos, al momento de insertar un nuevo dato o actualizar un dato existente, el balanceador de carga envía la orden a un microservicio cualquiera dependiendo del algoritmo de balanceo, en donde el microservicio insertará el dato en la base de datos y lo indexará localmente. El problema aquí es que los índices almacenados como archivos en una réplica pueden no ser exactamente iguales a los índices almacenados en otra ya que cada réplica es una

instancia independiente que se ejecuta en su propio contenedor con su propio volumen de datos e incluso en su propio host virtual o físico.

Es decir, sea MSI un microservicio que realiza búsquedas indexadas de tipo *full text* y MSI_k una instancia o réplica de MSI ejecutándose en un contenedor C_k en donde $k \in \{1..K\}$ siendo K el número total de réplicas y sea X_k el conjunto de datos indexados de la base de datos B en MSI_k (Figura 4). Suponiendo que en un instante de tiempo t_0 , $X_1=X_2=...=X_k=...=X_K$, en t_1 se envía una petición de inserción de datos al balanceador que la redirige a MSI_k el cual actualiza B y X_k . Al ser los índices independientes, en cualquier instante de tiempo t_i donde $i > 1$, $X_j \neq X_k$ donde $j \in \{1..K\}$ y $j \neq k$, de igual forma, la única instancia que mantiene coherencia entre la base de datos B y la indexación es MSI_k .

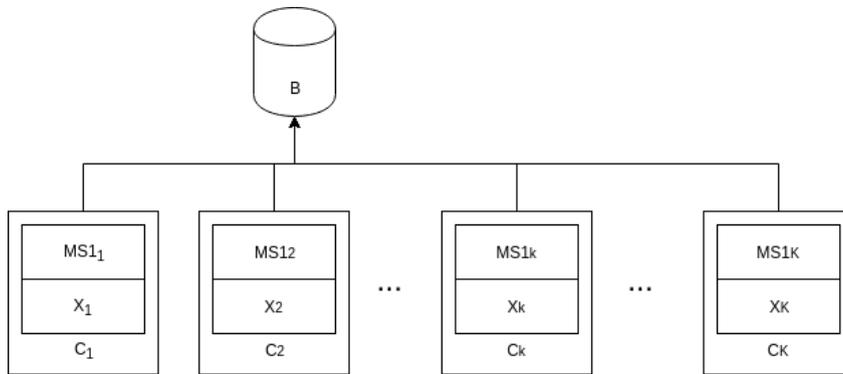


Fig. 4. Arquitectura de indexación no centralizada

Al momento de hacer una búsqueda del nuevo dato, este se encontrará si y solo si es buscado por el mismo micro servicio que lo insertó es decir MSI_k , pero al estar bajo el balanceo de carga, esto no se puede garantizar, por lo que, la probabilidad de encontrar el nuevo dato reducirá con cada nueva instancia de MSI que se ejecute.

Para solucionar el problema descrito anteriormente se utilizó Elasticsearch como administrador de índices, en donde toda la información de los índices se guarda en un solo lugar (indexación centralizada), así, no importa que microservicio ejecute el *insert* o el que busque el dato, los índices siempre estarán actualizados.

En la Figura 5 se muestra la arquitectura de la indexación centralizada utilizando Elasticsearch

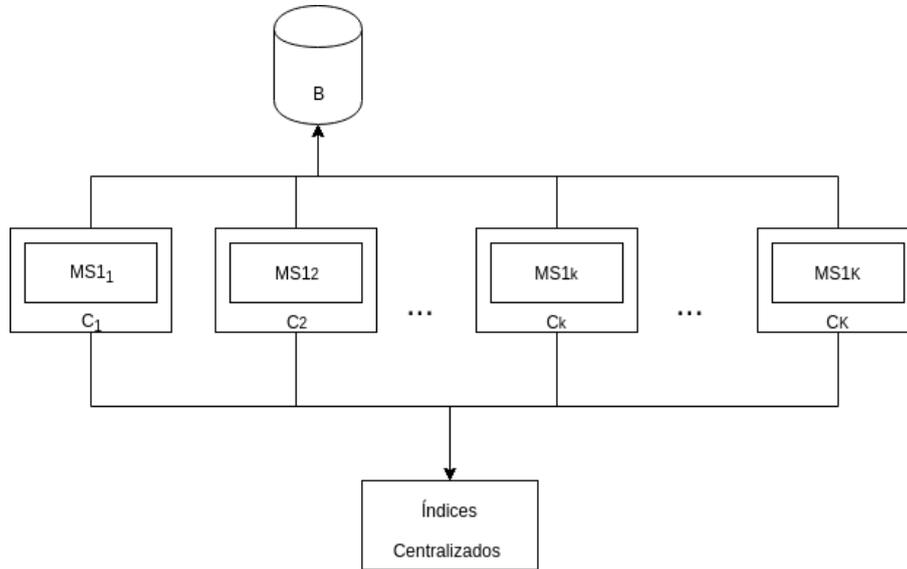


Fig. 5. Arquitectura de índices centralizados

Con la indexación centralizada se logra que para todo $MS1_k$ exista un único repositorio de índices X y que en cualquier instante de tiempo t_i en donde $i \in \{0..I\}$ siendo I el tiempo total en el que al menos una instancia de MSI se esté ejecutando, exista siempre coherencia entre X y B .

Caché centralizada

El manejo de la caché en cada réplica de los microservicios puede tener un problema similar al de la indexación. Supongamos que H_k es una caché de estructuras de datos o de resultados de llamadas a una función en el microservicio $MS1_k$, si en un instante de tiempo t_0 , $H1=H2=\dots=H_k=\dots=HK$, en t_1 el balanceador de carga recibe una petición de actualización de un dato en la base de datos B , la redirige a $MS1_k$, este actualizará la base de datos B , la caché H_k pero no las demás cachés de los microservicios réplica por lo que en cualquier instante de tiempo t_i donde $i>1$, $H_j \neq H_k$ donde $j \in \{1..K\}$ y $j \neq k$.

Si se realiza más de una consulta en cualquier instante de tiempo t_i donde $i>1$ dependiendo de la réplica que ejecute la petición, puede que el usuario final vea resultados distintos para la misma llamada.

A diferencia de la indexación, este problema sólo ocurre cuando se actualiza un dato y no cuando se inserta uno nuevo, puesto que, si el dato consultado no existe en la caché del microservicio que ejecuta la petición, éste lo buscará en la base de datos, lo registrará en su caché y devolverá el resultado correcto, por lo que eventualmente todas las réplicas tendrán el nuevo dato registrado en su caché.

El manejo de una caché transversal (Figura 6) para todos los micro servicios réplica permite que, en cualquier instante de tiempo, todas las réplicas $MS1_k$ accedan a una única caché por lo tanto siempre devolverán la misma respuesta a una petición

específica.

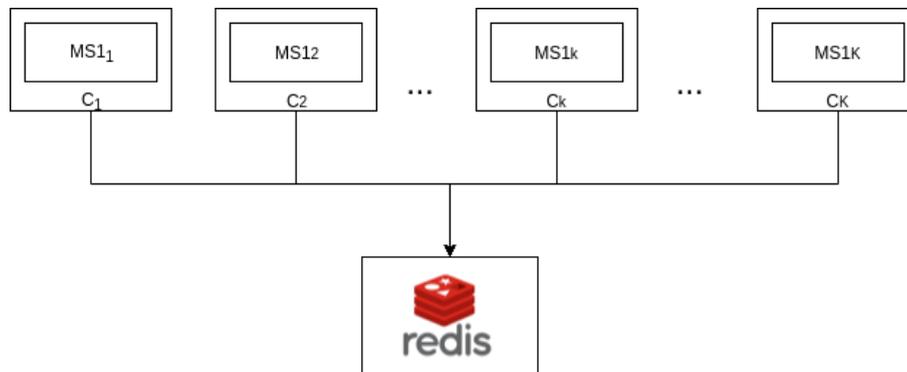


Fig. 6. Arquitectura de caché centralizada

En esta arquitectura se aprovechan los beneficios de la tecnología de Redis para implementar un clúster de cachés obteniendo un mejor rendimiento y alta disponibilidad.

Eventos y logs centralizados

Un aspecto importante en todos los sistemas, independiente de la arquitectura en la que estén implementados, es que tengan un adecuado manejo de sus eventos o logs. En los sistemas monolíticos, por lo general el servidor de aplicaciones almacena los logs del sistema en una sola ubicación, ya sea este un archivo o conjunto de archivos. De esta forma es relativamente sencillo monitorear el correcto funcionamiento de todos los componentes. En la arquitectura de microservicios y especialmente en los ambientes de balanceo de carga en el que las réplicas de un microservicio pueden ejecutarse en diferentes hosts sean estos virtualizados o físicos e inclusive en diferentes localidades geográficas, es importante tener un repositorio centralizado que almacene los logs de todos los microservicios y mediante el cual se pueda dar seguimiento desde un solo punto a todo el sistema.

Para esto en la arquitectura propuesta se ha implementado Graylog como repositorio de logs centralizados. Todas las réplicas de un microservicio e inclusive los diferentes microservicios o aplicativos que interactúan en el sistema envían sus logs a un mismo repositorio, Graylog almacena los logs en un repositorio MongoDB y utiliza Elasticsearch como motor de búsqueda, de esta forma mediante consultas sencillas se puede organizar los logs por microservicio o aplicativo e inclusive por palabras clave logrando así una correcta organización y manejo adecuado de los logs de cada componente del sistema.

Conclusiones y trabajos futuros

El uso de microservicios en un ambiente escalable no puede darse si no se tiene una adecuada orquestación de todos los microservicios y componentes involucrados. Por tal motivo, se ha planeado un caso práctico de implementación de una arquitectura de alta disponibilidad para microservicios en donde se ha podido constatar que una adecuada configuración permite que el balanceo de carga funcione eficientemente, logrando indexar y colocar en caché la información además de obtener un monitoreo completo de cada una de las instancias de los microservicios. La arquitectura propuesta puede ser fácilmente replicable en otras instituciones de educación superior que decidan transformar sus aplicaciones monolíticas a aplicaciones basadas en microservicios. Los resultados obtenidos de los microservicios puestos en producción en la Universidad de Cuenca han permitido poseer una nueva visión de los desarrollos actuales dentro de la Ingeniería de Software.

Los trabajos futuros para mejorar la arquitectura planteada serán:

- Implementación de un cluster de base de datos.
- Implementación de una herramienta para corto circuito de aplicaciones.
- Automatización del despliegue en caliente de instancias de microservicios de acuerdo al monitoreo que se realice.

Referencias

- 1 Francesco, Paolo Di.: Architecting Microservices. En: 2017 IEEE International Conference on Software Architecture Workshops (ICSAW), pp. 224 a 229. Gothenburg, Sweden: IEEE (2017)
- 2 Bakshi, K.: Microservices-based software architecture and approaches. En 2017 IEEE Aerospace Conference, pp. 1 a 8. Big Sky, MT, USA: IEEE (2017)
- 3 Dragoni, N., Giallorenzo, S., Lafuente, A. L., Mazzara, M., Montesi, F., Mustafin, R., & Safina, L.: Microservices: Yesterday, Today, and Tomorrow. En M. Mazzara & B. Meyer (eds.), Present and Ulterior Software Engineering, pp. 195-216 (2017).
- 4 Namiot, D., & Sneps-Sneppe, M.: On micro-services architecture. International Journal of Open Information Technologies, pp. 24 a 27 (2014)
- 5 Larrucea, X., Santamaria, I., Colomo-Palacios, R., y Ebert, C.: Microservices. IEEE Software 35, no 3, pp. 96 a 100 (2018)
- 6 Balalaie, A., Heydarnoori, A., & Jamshidi, P.: Migrating to Cloud-Native Architectures Using Microservices: An Experience Report. En A. Celesti & P. Leitner (eds.), Advances in Service-Oriented and Cloud Computing, vol. 567, pp. 201 a 215 (2016)
- 7 Kang, H., Le, M., & Tao, S.: Container and Microservice Driven Design for Cloud Infrastructure DevOps. 2016 IEEE International Conference on Cloud Engineering (IC2E), pp. 202 a 211(2016)
- 8 Yang, B., Sailer, A., Jain, S., Tomala-Reyes, A. E., Singh, M., & Ramnath, A. Service Discovery Based Blue-Green Deployment Technique in Cloud Native

- Environments. 2018 IEEE International Conference on Services Computing (SCC), pp. 185 a 192. (2018).
- 9 Chen, L.: Microservices: Architecting for Continuous Delivery and DevOps. 2018 IEEE International Conference on Software Architecture (ICSA), pp. 39 a 397. (2018).
 - 10 Kaur, S., Kumar, K., Singh, J., & Navtej Singh Ghumman: Round-robin based load balancing in Software Defined Networking. 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom), pp. 2136 a 2139 (2015).
 - 11 Bernard, E., Ferentschik, H., Fernandes, G., Grinovero, S., Memon, N. A., Morling, G., ... Smet, G.: Hibernate Search - Apache Lucene™ Integration. 206.
 - 12 Carlson, J. L., & Sanfilippo, S: Redis in action. Shelter Island, NY: Manning (2013).
 - 13 Al-Dhuraibi, Y., Paraiso, F., Djarallah, N., & Merle, P.: Autonomic Vertical Elasticity of Docker Containers with ELASTICDOCKER. 2017 IEEE 10th International Conference on Cloud Computing (CLOUD), pp. 472-479 (2017)
 - 14 Balalaie, A., Heydarnoori, A., & Jamshidi, P.: Microservices Architecture Enables DevOps: Migration to a Cloud-Native Architecture. IEEE Software, 33(3), pp. 42 a 52. (2016)
 - 15 M. L. Abbott y M. T. Fisher, The art of scalability: scalable web architecture, processes, and organizations for the modern enterprise, Second edition. New York: Addison-Wesley (2015)

Sistema de Recomendación en la Nube para Apoyar la Colaboración entre Investigadores

Fernando Barraza¹, Jose Luis Jurado²

¹ Cll 18 # 118-250, Cali Colombia
Universidad Javeriana

fbarraza@javerianacali.edu.co

² Cra 122 # 6-65, Cali Colombia
Universidad de San Buenaventura
jjurado@usbcali.edu.co

Resumen. Este artículo presenta un sistema de recomendación que apoya los procesos de colaboración entre investigadores. El sistema de recomendación utiliza un filtrado colaborativo sobre datos recopilados desde una plataforma de gestión de proyectos de investigación. Para realizar las recomendaciones utilizamos en un algoritmo de k-vecinos más cercanos en tres casos de uso diferentes. En el presente artículo presentamos el diseño del modelo, los algoritmos para cada caso de uso, el resultado de la evaluación de las recomendaciones y la arquitectura para la implementación del sistema en un servicio en la nube.

Palabras Clave: Colaboración, Aprendizaje Automático, Sistemas de Recomendación, Servicios en la Nube

Eje temático: Smartcampus y tendencias tecnológicas emergentes en la Universidad (aplicaciones de la inteligencia artificial, blockchain, analítica, IoT, impresión 3D, experiencia inmersiva).

1. Introducción

Actualmente, existen en Internet un conjunto de plataformas orientadas a la investigación como researchgate.com, mendeley.com, vivoweb.org, las cuales ofrecen a sus usuarios (investigadores) servicios que facilitan la identificación de oportunidades de colaboración con otros usuarios [1]. A pesar de la disponibilidad de tales plataformas, persisten aún limitaciones para que los investigadores puedan identificar y establecer vínculos de colaboración con sus colegas [2]. En el contexto latinoamericano, estas limitaciones fueron ilustradas y analizadas por Arcila et al [3] quienes encontraron un porcentaje menor de investigadores que utilizan plataformas de colaboración científica en comparación con Europa, Asia y América del Norte. De otra parte, no es común encontrar en las instituciones de educación superior en Latinoamérica, sistemas de información orientados exclusivamente a apoyar la gestión de proyectos y productos de investigación dejando el cubrimiento de dicha necesidad a la utilización que sus investigadores hagan de plataformas como las mencionadas anteriormente. En términos generales, el alcance de este tipo de sistemas de información es permitir el ingreso de la información relacionada con los proyectos de investigación, desde su estructura y contenido hasta los productos resultantes de dichos proyectos. Aún así con el panorama descrito anteriormente en relación a la escasa presencia y uso de plataformas y sistemas de apoyo a la investigación científica, las instituciones cuentan con la información de cientos de proyectos, ya sea en bases de datos o archivos de Word y Excel. Dicha información contiene datos que permiten caracterizar los proyectos y sus autores en relación a un amplio número de atributos como por ejemplo el tipo de investigación y las áreas de la ciencia que son de su interés. La disponibilidad de tal información, es la base para pensar que es posible implementar un sistema de recomendación cuyo objetivo sea mejorar la identificación de oportunidades de colaboración entre investigadores entre una misma o diferentes instituciones.

Con el objetivo anteriormente establecido, diseñamos un modelo de aprendizaje automático que permite implementar un sistema de recomendación (RS)²⁶ de apoyo a la colaboración entre investigadores. Los sistemas de recomendación son sistemas que predicen la calificación que un usuario otorgaría a un elemento (como libros, películas o personas) según las preferencias del usuario [4]. Sobre la base de las características de los investigadores y sus proyectos, realizamos una prueba de concepto que sugiere oportunidades de colaboración para los investigadores dentro de sus universidades o con otras instituciones. Probamos tres casos de uso para esta prueba de concepto. En este documento, presentamos el diseño del sistema de recomendaciones, los algoritmos para los casos de uso seleccionados, los resultados en términos de la precisión de las recomendaciones y la implementación del sistema en un servicio en la nube.

2. Motivación

Los sistemas de recomendación ahora se utilizan ampliamente para mejorar la

²⁶ Que de ahora en adelante nos referiremos como RS

experiencia del usuario con los datos recopilados de la interacción entre los usuarios y los sistemas de software, en particular en aplicaciones web como portales, música en línea y sitios de películas en línea. Hay muchos ejemplos de cómo usar RS como Amazon.com, Netflix.com entre otros. Básicamente, son dos los tipos de técnicas de RS más utilizados, *Filtrado Colaborativo* y *Basado en Contenido*, (CF y CB, por siglas en inglés). El primero intenta hacer sugerencias o predicciones basadas en la similitud de los usuarios o elementos y el segundo se basa en la similitud de una descripción del contenido de los elementos. En terminología de RS, un artículo es una cosa que podría ser calificada por cualquier usuario y un usuario es la persona que califica un artículo según su preferencia hacia el artículo. Sin embargo, en cualquier modelo, la idea detrás de la RS es intentar hacer recomendaciones de artículos o usuarios en función de qué tan similares sean a un artículo o usuario específico, respectivamente. Hay muchos algoritmos para calcular la similitud entre usuarios, elementos o ambos (correlación de *Pearson*, distancia euclidiana, similitud de *Jaccard*, etc.) [5]. Aunque un RS funciona bien en la mayoría de los casos, hay situaciones en las que calcular la distancia entre elementos no es suficiente. Por ejemplo, un RS no funciona bien si hay muchos usuarios que no ven la mayoría de los elementos. En esos casos se hace necesario mejorar el RS utilizando otros enfoques como los modelos de clasificación [6]. Un enfoque posible es utilizar algoritmos de *Aprendizaje Automático* (*Machine Learning*, en inglés) que permiten calcular un valor de una variable objetivo de acuerdo a un modelo de clasificación que ha sido entrenado previamente. De esta forma se puede construir un *Ranking* con todas las predicciones hechas a todos los artículos para un usuario de interés específico. El RS mostrará entonces al usuario los resultados, con un criterio de recomendar de forma predeterminada los n -primeros artículos del ranking obtenido.

La Colaboración en Procesos de Investigación

El concepto de trabajo colaborativo, se debe entender inicialmente como ese ejercicio de múltiples individuos trabajando juntos de una manera planificada en un mismo proceso de producción o en procesos de producción diferentes pero conectados. En este sentido los procesos de investigación en su naturaleza deben verse como actividades colaborativas donde la coordinación y colaboración se interrelacionan y se complementan. Mientras que la coordinación implica la presencia de una entidad externa que controla la interacción entre los miembros del sistema, la colaboración tiene un enfoque más orientado a que el control de dicha interacción se lleva a cabo por los mismos miembros del sistema, de una manera distribuida, como lo afirma [10]. Lo que se busca finalmente es que los procesos de investigación se beneficien de las sinergias derivadas por la cooperación de trabajos realizados en conjunto por diferentes investigadores que permite a través de la colaboración entre estos, la generación de artefactos y productos de alto impacto para la investigación. De lo anterior podemos inferir que, si la colaboración es efectiva, los procesos de investigación mejorarán en sus resultados con beneficios a una comunidad educativa y de investigación.

No obstante, aún teniendo claridad sobre la importancia de la colaboración en investigación, es importante también reconocer cuales de las capacidades de investigación de una universidad son susceptibles de mejorar incrementando la colaboración en sus procesos. La figura 1 muestra un modelo de capacidad basado en

el concepto de Arquitectura Empresarial para una institución de educación superior en el Reino Unido.

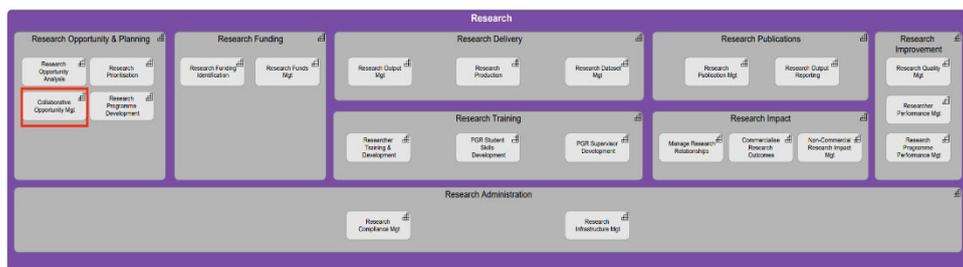


Fig 1. Modelo de Capacidad de los procesos de investigación de una Universidad en el Reino Unido (tomado de https://www.ucisa.ac.uk/representation/activities/cap_model)

Si bien es posible indicar que cualquiera de las capacidades de investigación son susceptibles de ser mejoradas introduciendo elementos de colaboración en sus procesos, nuestro trabajo se ha centrado específicamente en la capacidades de “Gestión de Oportunidades de Colaboración” (*Collaborative Opportunity Mgt*, encerrada en cuadro rojo en la figura 1). En consecuencia, el trabajo que estamos presentando en este documento se apoya en el uso de tecnologías innovadoras, como lo son las basadas en Aprendizaje Automático y Sistemas de Recomendación, para mejorar entonces la capacidad de gestionar por parte de los mismos investigadores las oportunidades de colaboración en investigación.

3. Diseño del Sistema de Recomendación

El objetivo principal de implementar un RS en la plataforma de investigación es propiciar la colaboración entre investigadores (usuarios en terminología de RS). Por lo general, los investigadores conocen un círculo limitado de colegas que trabajan en sus mismas áreas temáticas pero dicho conocimiento no llega hasta el nivel de detalle que les permita conocer cuales de los proyectos de sus colegas podrían tener una mayor relación con sus propios proyectos (elementos en terminología RS). En este sentido, un investigador podría resolver algunas inquietudes de forma sencilla, como por ejemplo ¿Quiénes son los investigadores que tienen intereses similares a los míos? pero le sería más complejo resolver dudas como ¿Qué proyectos (de mis colegas y otros) son similares a los míos? Más aún, conocer proyectos similares podría ser en algún momento de menos interés para un investigador si lo que busca es trabajar en proyectos interdisciplinarios, es decir que abarcan diferentes áreas temáticas de la ciencias. En ese caso una pregunta más relevante podría ser ¿A qué proyectos debería echar un vistazo? De otra parte, es muy común en los procesos de evaluación de propuestas de investigación, conocer quienes podrían ser asesores en la formulación de los proyectos o posibles evaluadores de los mismos. En este caso, una pregunta pertinente sería ¿Quién podría ser la persona indicada para realizar una revisión por pares de mi propuesta de proyecto? En general, así como ocurre en los procesos de evaluación de

propuestas de proyectos, hay muchas más preguntas pertinentes en otros de los procesos de investigación donde se presenta interacción entre investigadores. Nuestro argumento principal es que las anteriores preguntas pueden intentar ser resueltas con un modelo apropiado de aprendizaje automático que este basado en la información existente sobre los proyectos de los investigadores. En nuestro caso de estudio los proyectos de investigación están registrados en una plataforma de gestión de proyectos comercial disponible en la nube.

Teniendo en cuenta el contexto propio de nuestro problema, elegimos un modelo de RS de tipo Filtrado Colaborativo. Dicho modelo se basa en el supuesto de que los investigadores con proyectos similares tienen intereses similares. Definimos entonces el concepto de similitud entre dos proyectos, como la medida de la distancia entre ellos. Igualmente, definimos el concepto de similitud entre dos investigadores como la medida de la distancia entre ellos, es decir, qué tan similares son los proyectos de los dos investigadores.

Formalmente, sea R el conjunto de todos los investigadores y P el conjunto de todos los proyectos registrados en la plataforma de investigación, entonces $P_{ri} = \{p_{r1}, p_{r2}, \dots, p_{rn}\}$ es el subconjunto de proyectos que pertenecen al investigador r , donde $i = 1, 2, 3, \dots, n$.

Para medir la distancia entre dos proyectos utilizamos la métrica de similitud de coseno. La similitud de coseno es una medida que calcula el coseno del ángulo entre dos vectores. Aplicado en una comparación entre proyectos, cada vector es un proyecto que tiene como magnitud el recuento de palabras presente en el texto del documento del proyecto (*tf-idf*). La métrica del coseno nos da entonces una medida de cuánto se relacionan dos proyectos al observar el ángulo de dos vectores en un espacio normalizado, esto en lugar de un valor expresado simplemente como una magnitud. *tf-idf* es una técnica ampliamente utilizada en el campo de recuperación de información (IR) [7].

La similitud de coseno ω entre dos proyectos se define por,

$$\omega(\vec{p}_1, \vec{p}_2) = \cos \theta = \frac{\vec{p}_1 \cdot \vec{p}_2}{\|\vec{p}_1\| \|\vec{p}_2\|} \quad (1)$$

Ahora que sabemos cómo obtener la similitud entre dos proyectos, podemos calcular cuán similares son todos los proyectos de un investigador r con un solo proyecto x que pertenece a otro investigador q . Para hacer esto, utilizamos la fórmula que contiene el promedio de las distancias (1) de los pares de proyectos formados entre

los proyectos P_{ri} del investigador r con los proyectos p_{qx} del otro investigador q :

$$d(P_r, p_{qx}) = \frac{\sum_{i=1}^n \omega(p_{ri}, p_{qx})}{n} \quad (2)$$

Con la anterior ecuación, construimos un conjunto de entrenamiento que nos dice para cada investigador, la puntuación de similitud entre los proyectos propios con un

solo proyecto que pertenece a otro investigador. La puntuación es el resultado de la función d en (2). Llamamos a esta puntuación Similitud de Proyecto.

Tabla 1. Extracto del conjunto de datos de entrenamiento obtenido.

Investigador i	Nombre de Investigador	Identificador de Proyecto j	Similitud de Proyecto $d(R_i, P_j)$
1	Ana	6	0.875
2	Juan	7	0.478
3	Luis	545	0.277
4	Pedro	545	0.653
5	Martha	88	0.788

La tabla 1 muestra un ejemplo de cómo sería el conjunto de entrenamiento. En términos de Aprendizaje Automático, la puntuación es la clase de salida, y para el RS es la calificación predicha.

Uso del algoritmo k-nn para hacer las Recomendaciones

Como dijimos anteriormente, hay muchas preguntas que pueden resolverse con el modelo propuesto. En este trabajo, mostramos cómo realizar tres tipos de recomendaciones: La primera, es aconsejar a un investigador sobre quiénes son los otros investigadores con quienes dicho investigador podría buscar algún tipo de colaboración. El segundo es sugerirle qué proyectos de sus colegas son interesantes para él y que por lo tanto debería darles un vistazo. El último es, dado un proyecto de un investigador, que otros investigadores podrían ser sus asesores o revisores pares.

Para hacer esas recomendaciones, primero, definimos qué tan similares son los investigadores en función de sus proyectos. Hacemos esto comparando a cada investigador con cada uno de los otros en función del cálculo de la puntuación de similitud entre ellos. Estos cálculos forman el modelo entrenado. Luego usamos un algoritmo de vecinos más cercanos (k -nn) para predecir la puntuación de similitud [8]. Lo anterior se hace en dos pasos:

1. Examinar cada elemento (investigadores) en el modelo entrenado y calcular una puntuación de similitud desde ese elemento de datos a la muestra específica (investigador).
2. Clasificar la muestra (investigador) como la clase mayoritaria entre k muestras (investigadores) en el modelo entrenado que tenga una distancia mínima con el investigador (muestra).

Ahora que tenemos un modelo entrenado, podemos pasar a realizar las recomendaciones propuestas.

Encontrando investigadores similares

Como dijimos antes, hay algunos métodos para calcular una puntuación de similitud.

En nuestro trabajo, hemos elegido la distancia euclidiana para usarla en el paso 1 dentro del algoritmo k -nn. La figura 1 muestra una representación grafica de la similitud de dos investigadores, donde los proyectos son los ejes x y y , y los puntos en el espacio son los valores de similitud de cada investigador con los proyectos representados en los ejes.

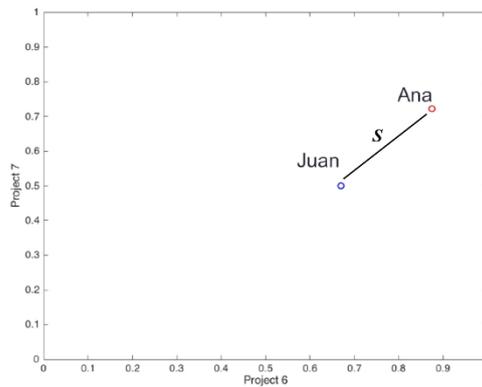


Fig. 2. Distancia Euclidiana entre dos investigadores

El concepto detrás de esto es que mientras más cerca estén dos puntos en el espacio, más similares serán los dos investigadores. El denominador de la función en la ecuación (3) da valores más altos para los investigadores que son más similares por lo que es necesario obtener en cambio el valor inverso de la función.

Por ejemplo, supongamos que r_1 representa la investigadora llamada Ana y que r_2 al investigador de nombre Juan. La puntuación de similitud S entre ellos para los proyectos 6 y 7 se calcula por medio de,

$$S(r_1, r_2) = \frac{1}{1 + \sqrt{(d(r_1, p_6) - d(r_2, p_6))^2 + (d(r_1, p_7) - d(r_2, p_7))^2}} \quad (3)$$

Debido a que el gráfico de la figura 1 que representa la similitud entre dos investigadores es bidimensional, solo podemos observar la similitud de dos proyectos a la vez, pero el principio es el mismo para conjuntos más grandes de puntajes de similitud. Con este principio en mente, podemos calcular el puntaje de similitud de un investigador con cada uno de los demás investigadores y luego clasificarlos para obtener el k mejor clasificado entre todos los investigadores. Por ejemplo, para obtener personas que sean más similares a Ana podemos usar la siguiente función,

$$reco_researchers(r_1) = TopMost(S(r_1, r_j), k) \quad (4)$$

donde $j = 1$ hasta el total de investigadores y k es cuántas recomendaciones queremos obtener.

Con lo anterior podemos entonces obtener la primera recomendación mencionada anteriormente.

Recomendando Proyectos

Para saber qué proyectos recomendar a un investigador, podemos mirar a la persona que más se parece a él y mirar entre sus proyectos. Sin embargo, este enfoque podría devolver proyectos que pertenecen a la persona más similar al investigador al que queremos dar la recomendación (investigador objetivo) pero que tienen un puntaje de similitud bajo con sus proyectos. Para resolver este problema, debemos obtener la puntuación de similitud de los proyectos a través de una puntuación que se pondere de acuerdo a la similitud entre los investigadores. Tomamos entonces el valor de similitud que se obtiene entre cada par de proyectos y lo multiplicamos por el valor de similitud que se tiene entre los investigadores de dichos proyectos.

Por ejemplo, si queremos saber qué proyecto recomendar a Ana, podemos usar la siguiente ecuación:

$$reco.proy(r_1) = TopM \text{ost}(\hat{S}(r_1, r_j) \square d(P_j, p_n), k) \quad (5)$$

donde $j = 1$ hasta el total de investigadores, $n = 1$ hasta total de proyectos, y k es cuántos proyectos queremos obtener en la recomendación.

Con la formula anterior se resuelve entonces la segunda recomendación propuesta.

Recomendando Pares Evaluadores

La idea de esta recomendación es dar a los nombre de k investigadores que podrían ser los revisores de un proyecto determinado. Para resolver esto, usamos un enfoque similar al anterior: Tomamos el puntaje de similitud entre cada proyecto con el proyecto objetivo y lo multiplicamos por el puntaje de similitud obtenido entre cada investigador con el proyecto objetivo. Siguiendo con el ejemplo de Ana, si queremos dar recomendaciones para el proyecto con *id* igual a 6 podemos usar la siguiente formula,

$$reco.peers(P_6) = TopM \text{ost}(\hat{X} \square d(P_j, p_6), k) \quad (6)$$

donde $j = 1$ hasta total el total de investigadores, $n = 1$ hasta el total de proyectos y k es el número de revisores que se desean obtener. Es importante anotar que debemos omitir a Ana en el conteo de j porque si no ella obviamente será la primera en la lista de personas a ser recomendadas.

Con la ecuación (6) hemos resuelto entonces la tercera recomendación.

Evaluando las recomendaciones

Con respecto a la evaluación de RS, hay muchas propiedades que deben considerarse como precisión de predicción, cobertura, novedad, casualidad, entre muchas otras [9]. En nuestro trabajo, elegimos como medida de evaluación la precisión en la predicción. Para obtener una medida de esta propiedad hay algunas métricas que pueden ser utilizadas. Estas son calcular la precisión de las predicciones de calificación, medir la precisión de las predicciones de uso y medir la precisión de las clasificaciones de los elementos. Consideramos que la precisión es la medida más adecuada para nuestro RS debido a la importancia del orden en que se presentan los proyectos e investigadores a los usuarios en las plataformas de investigación.

Para calcular esta medida usamos una métrica de clasificación basada en la utilidad. Con esta métrica, obtenemos una cantidad que nos indica que tan útil es la lista de recomendaciones presentadas al usuario de la plataforma. Es importante anotar que la utilidad de una lista de recomendaciones es aditiva y está dada por la suma de las utilidades de las recomendaciones individuales. Al mismo tiempo, la utilidad de cada recomendación es la utilidad del artículo recomendado, descontado por un factor que depende de su posición en la lista de recomendaciones. Este último hecho se debe a que se supone que, a pesar de que los usuarios miran la lista desde el principio hasta el final, como en los motores de búsqueda, la probabilidad de que los usuarios miren los elementos finales es menor que los primeros elementos.

La ecuación para calcular la utilidad se da en (7). Aplicamos esta ecuación a los tres casos de uso de recomendaciones cubiertos inicialmente por nuestro RS.

$$Rating_u = \sum_j \frac{\max(r_{ui} - d, 0)}{2^{\frac{j-1}{\alpha-1}}} \quad (7)$$

donde i_j es el elemento en la posición j -ésima, r_{ui} es la similitud del usuario u del elemento i , d es una calificación neutral dependiente de la tarea, y α es un parámetro para controlar la rapidez con la que el valor de las posiciones disminuye en la lista clasificada.

Digamos que, dependiendo del caso de uso de recomendación que estemos evaluando, el elemento serán los proyectos o los investigadores, por lo que la ecuación en (7) debe ajustarse con los valores apropiados.

La tabla 2 muestra los resultados obtenidos para las pruebas del RS en cada caso de uso. Se puede analizar a partir de dichos resultados que el RS se comportó de manera aceptable, con un valor de aproximadamente el 65% en promedio para todas las recomendaciones hechas.

Tabla 2. Resultados de la evaluación del RS de acuerdo a la precisión de las recomendaciones

Caso de Uso	Numero de Recomendaciones Hechas	Rating Promedio
Investigadores con Intereses Similares a un Investigador	92	77%
Proyectos de Interés para un Investigador	70	71%
Pares Evaluadores de un Proyecto	55	48%

El numero de recomendaciones hechas dependió de los resultados obtenidos que fueron considerados como exitosos, de un total de 100 casos totales disponibles y ejecutados. De la tabla anterior se puede observar que, teniendo en cuenta que la formula en el segundo caso de uso se apoya en el primero, el RS se comportó mejor y por encima del 50% en los casos donde se buscaban obtener recomendaciones basadas en la similitud entre investigadores y menor al 50% en el caso donde se buscaban recomendaciones basadas en la similitud entre proyectos. Consideramos que el resultado obtenido es más que aceptable para el primero modelo del RS que se ha diseñado. Las mejoras en los resultados en la precisión del RS dependerá en gran medida en las mejoras al modelo utilizado, incluyendo otros atributos relacionadas con los proyectos o los mismos investigadores.

Implementación del RS en un servicio en la nube

El objetivo final del proyecto era el poner en operación el RS en un entorno computacional (llamado también *ambiente de producción*) donde se ejecutarán en tiempo real los cálculos necesarios para brindar las recomendaciones in vivo a los usuarios de la plataforma de investigación. En las secciones anteriores hemos recorrido a través del diseño del modelo, las actividades relacionadas con el funcionamiento de un RS. Estas actividades representan en buena parte el flujo de trabajo de un RS, el cual es una vez desplegado en producción se comporta de manera cíclica. La figura 2 muestra un ciclo de trabajo típico para un RS.

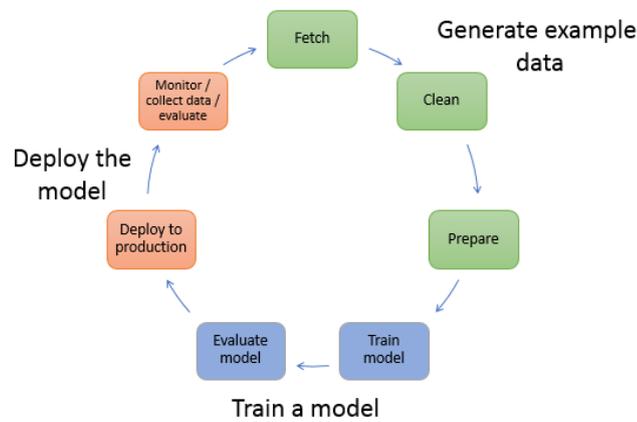


Fig 3. Ciclo típico de un RS (tomado de AWS)

De la figura anterior se pueden observar tres fases principales (diferencias en tres colores diferentes): Generar los datos de ejemplo, Entrenar el modelo y Desplegar el modelo en un ambiente de producción. Las dos primeras fases se realizaron a través de las pruebas de concepto sobre los tres casos de uso mencionados anteriormente. Una vez las tres pruebas de concepto fueron satisfactorias, se procedió a desplegar el RS como un servicio en la nube. El despliegue en la nube de un sistema en producción como el RS implica suscribir, configurar y poner en marcha un conjunto de recursos computacionales que sigue un patrón predeterminado de arquitectura en la nube. En nuestro caso utilizamos la nube del proveedor AWS²⁷ el cual ofrece el servicio *Sagemaker*, el cual es un servicio especializado para implementar modelos de Aprendizaje Automático bajo diferentes algoritmos, como lo es para nuestro caso el algoritmo *k-nn*. La figura 2 presenta la arquitectura del RS implementado en AWS, para la cual seguimos las recomendaciones y buenas prácticas ofrecidas por este proveedor.

²⁷ <http://www.aws.com>

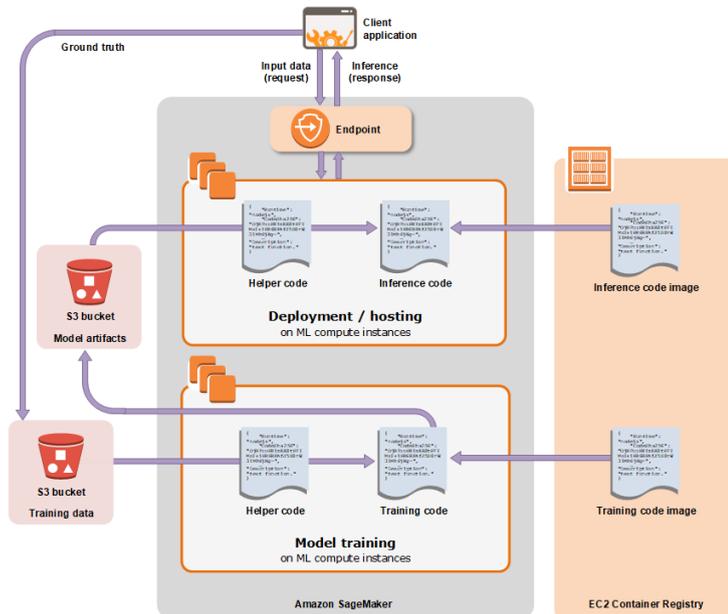


Fig 4. Modelo de despliegue del Servicio RS en la nube (tomado de AWS)

En la figura anterior se pueden identificar diferentes componentes de la arquitectura:

- Un servicio de Sagemaker el cual contiene dos instancias de cómputo: Una instancia ML para entrenamiento en la cual se ejecutan los cálculos para construir el modelo entrenado y otra instancia ML para inferencia donde, basados en el modelo entrenado, se ejecutan los cálculos necesarios para dar las recomendaciones solicitadas por los programas clientes del RS.
- Un conjunto de datos de entrenamiento: Estos datos se obtienen desde la base de datos relacional de la plataforma de investigación. Los datos se encuentran en formato *csv* siguiendo un estilo similar a la tabla 1 y son almacenados en un *bucket S3* de AWS.
- Un conjunto de artefactos del modelo: Estos artefactos corresponde a los datos del modelo ya entrenado. En nuestro caso son archivos en formato JSON los cuales son consumidos por la instancia donde se hacen las inferencias correspondientes a las recomendaciones solicitadas al RS. Un contenedor de registro: En este componente se almacenan el código que implementa los algoritmos *k-nn* del RS. En nuestro caso utilizamos el lenguaje *python* para la implementación de los tres casos de uso que comprenden el alcance del RS. El código catalogado como de ayuda (*helper*) es el código de las librerías propias de python que hacen parte de la implementación del RS.
- Una aplicación cliente: Este componente de tipo programa tiene dos objetivos. Por un lado ofrece un *front-end* el cual permite a un usuario final de la plataforma (investigador) obtener las recomendaciones ofrecidas por el RS. Para esto, el programa envía un *request* al servicio de Sagemaker con los parámetros propios de la recomendación solicitada. El segundo

objetivo es el de suplir al RS de los datos necesarios para realizar el entrenamiento del modelo o *Ground Truth*. En la arquitectura este término se refiere al proceso de recopilación de los datos objetivos apropiados (comprobables) que harán parte del entrenamiento del modelo. En nuestro caso dichos datos son obtenidos desde la base de datos relacional de la plataforma de investigación, que es la aplicación donde los usuarios registran sus proyectos.

- Un *end-point* de Sagemaker. Es el punto de interfaz donde un programa cliente puede hacer solicitudes de recomendación al RS. Para esto la aplicación cliente debe pasar en su solicitud el parámetro adecuado para cada caso de uso que hemos implementado. Por ejemplo para recomendar proyectos de interés de un investigador (caso de uso numero 2) deberá pasarse como parámetro el identificador del investigador objetivo.

La arquitectura en la nube desplegada para el RS ha mostrado tiempos de respuesta satisfactorios para conjuntos de datos relacionados con más de 10.000 proyectos de investigación tomados como datos para entrenamiento del modelo. Igualmente la característica de arquitectura elástica ofrecida por el proveedor del servicio en la nube garantiza la posibilidad de escalar el servicio a un número mayor de conjuntos de datos de entrenamiento, lo que posibilita no solo pensar en la implementación de más casos de uso de recomendación si no en la expandir el alcance del RS a más instituciones.

4. Conclusiones y trabajo futuro

Hemos presentado el diseño de un Sistema de Recomendación que ofrece recomendaciones sobre investigadores similares, proyectos de interés y posibles revisores. En el diseño del RS hemos presentado cómo construir el conjunto de datos de entrenamiento y las fórmulas para implementar los algoritmos de recomendaciones para una primera prueba de concepto. Además, presentamos una métrica inicial para la evaluación de RS teniendo en cuenta la utilidad de las recomendaciones para los usuarios de la plataforma. Las primeras evaluaciones sugieren un nivel de precisión aceptable para los tres casos de uso de las recomendaciones a obtener. Igualmente hemos presentado la arquitectura de despliegue del RS en un servicio en la nube.

Este trabajo permite evidenciar indicios satisfactorios, en las redes de colaboración donde el uso de algoritmos k-vecinos más cercanos generan interacciones precisas y lógicas, las cuales ayudan al acoplamiento de tareas, intereses comunes e identificación de perfiles con similitudes compartidas en procesos y proyectos de investigación. Los modelos de recomendación que se logran identificar y sus posteriores resultados, generan una atractiva prospectiva en las tareas de cooperación entre individuos y la capacidad de generar conocimiento compartidos entre pares. Todo esto contribuye en el mejoramiento de las capacidades de gestión en investigación de los propios investigadores y por ende de la institución.

En un futuro cercano, tenemos en mente ampliar el alcance de la implementación del RS en otras de las capacidades de investigación de las instituciones de educación superior. Igualmente esperamos mejorar el rendimiento del RS implementando mediciones de evaluación en línea para comparar si las recomendaciones son realmente

útiles para los investigadores. Además, esperamos mejorar el modelo con la inclusión de otras características como el año del proyecto, las instituciones involucradas, así como otra información de relevancia del proyecto.

Al final, esperamos que nuestro trabajo promueva el uso de técnicas innovadoras, como lo son las de aprendizaje automático, para mejorar los procesos de investigación en las instituciones de educación.

Referencias

1. Bullinger, A., Hallerstedde, S.H., Renken, U. Soeldner, JH. Proceedings of the Sixteenth Americas Conference on Information Systems, Lima, Peru, August 12-15, (2010)
2. Lee, S., Bozeman, B. The Impact of Research Collaboration on Scientific Productivity. *Social Studies of Science*, Vol. 35, No. 5, Scientific Collaboration, pp. 673-702. (2005)
3. C. Arcila, J.L. Piñuel y M. Calderón. La e-investigación de la comunicación: Actitudes, herramientas y prácticas en los investigadores iberoamericanos. *Revista Comunicar*, no 40, v. XX, *Revista Científica de Educomunicación*; ISSN: 1134-3478; páginas 111-118. (2013)
4. Sharma, L. Gera, A. A Survey of Recommendation System: Research Challenges. *International Journal of Engineering Trends and Technology (IJETT) - Volume 4, Issue5* (2013)
5. T. Segaran, *Programming Collective Intelligence*, First ed., Ed. O'Reilly Media, pages. 117-127. (2007)
6. S. Owen, R. Anil, T. Dunning, E. Friedman. *Mahout in Action*. Manning Publications. ISBN: 9781935182689. (2015)
7. C. Manning, P. Raghavan, H. Schütze. *Introduction to Information Retrieval* 1st Edition. Ed. Cambridge University Press. ISBN-10: 0521865719. (2014)
8. S. B. Kotsiantis. Supervised Machine Learning: A Review of Classification Techniques. *Proceedings of the 2007 conference on Emerging Artificial Intelligence Applications in Computer Engineering: Real Word AI Systems with Applications in eHealth, HCI, Information Retrieval and Pervasive Technologies*. IOS Press Amsterdam, The Netherlands. ISBN: 978-1-58603-780-2. Pages 3-24 (2008).
9. G. Shani, A. Gunawardana. Evaluating Recommendation Systems. Chapter 8 of the book *Recommender Systems Handbook*. Ed. Springer. ISBN: 978-0-387-85819-7. (2016)
10. Garousi, V., Petersen, K., & Ozkan, B. (2016). Challenges and best practices in industry-academia collaborations in software engineering: A systematic literature review. *Information and Software Technology*, 79, 106-127.

Capítulo 7

Tecnologías en la Innovación - Investigación

Experiencia de la Universidad del Rosario en la Implementación de Servicios y Plataformas para Computación de Alto Desempeño y de Apoyo a la Investigación

Rodolfo Arturo Cáliz Ospino, Consuelo Ardila Castro,
Fredy Harvey Sáenz Guzmán

Universidad del Rosario
Tecnología Informática y Comunicaciones
Av. Jiménez # 4 – 09, Bogotá, Colombia
Rodolfo.caliz@urosario.edu.co
consuelo.ardila@urosario.edu.co, fredy.saenz@urosario.edu.co

Resumen. Este documento se centra en el diseño y en la evaluación de sistemas audiovisuales telemáticos, los cuales permiten a los intérpretes experimentar sonidos, imágenes, expresiones colaborativas y colectivas en ámbitos telemáticos. Basándose en enfoques de interacción humano-computador (HCI) y métodos de usabilidad de tercera ola, en los que se abordan elementos que profundizan más en cómo las personas se relacionan gracias al uso de tecnologías y menos en cómo el usuario cumple sus tareas, se lleva a cabo una serie de estudios de usuarios en diferentes prototipos planteados y desarrollados para tal fin. También se estudian algunas participaciones *performáticas* distribuidas con el objeto de encontrar puntos de inflexión y características útiles que ayuden a las propuestas de prototipos previos así como la construcción de un prototipo impulsado por técnicas de colaboración y evaluación, eje central de la investigación que da pie a este documento, la cual busca diseñar un espacio de comunicación colectiva entre intérpretes telemáticos.

1. Introducción

Según MacDonald y Atwood (MacDonald & Atwood, 2013) “la evaluación ha sido un tema dominante en HCI durante décadas, pero está lejos de ser un problema resuelto”. Esto es particularmente evidente en el contexto de los *performances* telemáticos y del desarrollo de interfaces de aplicaciones sonoras e interpretativas en vivo, donde la experiencia del usuario con interfaces lúdicas o creativas es a menudo marcado por una calidad idiosincrásica (Jordà, 2002). Como resultado, las peculiaridades y especificidades en la evaluación de tales interfaces quizá sea un fenómeno inevitable, haciendo que la búsqueda de una solución de “talla única” sea potencialmente inútil. En cambio, se propone que los desarrolladores investiguen la posibilidad de adaptar las técnicas existentes en el diseño de interfaces, un enfoque también defendido por Kiefer, Collins y Fitzpatrick (2008) o diseñar nuevas, si es necesario. La selección de técnicas existentes, como las descritas a lo largo de este documento, puede a su vez estar motivada por varios factores, incluyendo la disponibilidad de las herramientas necesarias, el grado en el que el contexto deseado de una técnica coincide con el examinado y el nivel de modificación requerido para adaptar una técnica de un dominio de aplicación a otro. Adicionalmente, se contempla esta propuesta de estructura metodológica como un claro ejemplo del paradigma de la investigación mixta, ya que se complementan los cuestionarios con discusiones cualitativas abiertas, datos cuantitativos registrados y entrevistas no estructuradas, aumentando así la profundidad de la retroalimentación que pudiéramos obtener de nuestros usuarios. Esto condujo, a su vez, a mejoras en el test y la evaluación con estos permitió reacondicionar y replantear la propuesta de los prototipos.

En los enfoques clásicos del diseño centrado en el usuario, como la usabilidad, se suele poner un fuerte énfasis en el grado en que los usuarios pueden realizar con éxito tareas con un determinado sistema, modelo o prototipo bajo evaluación. Con un número cada vez mayor de disciplinas recurriendo a la investigación de HCI para orientar el diseño de sistemas no solo usables sino también expresivos o simplemente más atrayentes, muchos investigadores se enfrentan a las deficiencias de tal enfoque basado en tareas. Marti y Bannon (2009), por ejemplo, pide un mejor marco para conceptualizar las actividades humanas tanto interpersonales como de comportamiento. De manera similar, Kaye (2007) se pregunta: “¿dónde quedan las estructuras metódicas para el desarrollo de tecnologías planteadas, no para cumplir las tareas, sino para tener experiencias, para expresar la propia identidad, para coquetear y argumentar y vivir?”. A su vez, las insuficiencias del enfoque basado en tareas en el examen de los aspectos de *performance* más allá de la usabilidad, han llevado a la aparición de lo que ahora se conoce como “HCI de tercera onda o de tercer paradigma”, una tendencia descrita por Kiefer, Collins, & Fitzpatrick (2008).

Es así como la evolución de las formas en que la tecnología se utiliza y como la computación se incorpora cada vez más en la vida cotidiana, promueve un enfoque de experiencia en lugar de uno basado en tareas de diseño impulsado por el usuario. En este sentido, estimula lo que describen Fallman y Waterworth (2005) cuando hablan de un enfoque en “experiencias más que en desempeño; diversión y jugabilidad en lugar de tasa de error; y la sociabilidad y las cualidades afectivas más que la capacidad de aprendizaje”. Como resultado, el HCI de tercera ola es particularmente adecuado para el diseño y la evaluación de nuevas interfaces audiovisuales interactivas. Una visión

paralela pero dentro del contexto exclusivo de lo sonoro es presentada por Johnston (2011) y El-Shimy & Cooperstock (2016), quienes desde perspectivas similares postulan que “la evaluación es mejor considerada como un componente de un examen más amplio de la interfaz musical”.

A partir de estos análisis y propuestas, se planea el análisis con la obra telemática, titulada MirroR. Para este análisis se define la aplicación de un test multivariable (Andersen, O'Rourke, Liu, Snider, Lowdermilk, Truong, 2012) y entrevistas semi-estructurada (Wood, 1997), sobre un prototipo de carácter evolutivo (Padhye & Kalia, 2009), haciendo uso de un método orientado hacia el usuario pero aplicando características de experiencia más que de tareas (Doering, Pflieger, Kray, & Schmidt, 2010).

2. Procedimientos y diseños

El diseño de entornos de respuesta para el *performance* de interacción distribuida, adoptó un modelo de diseño que hibrida los enfoques descritos. El modelo inicia con una comprensión temprana del usuario objetivo, seguida de ciclos iterativos de pruebas formales, entrevistas semiestructuradas, mejoras en el diseño, ampliación de variables del prototipo, análisis y evaluaciones. La elección de esta metodología particular — impulsada por el usuario— fue motivada en parte por la amplia naturaleza de los objetivos iniciales de investigación.

Si bien se ha establecido la mejora del *performance* de interacción distribuida como el principal objetivo y se han creado directrices para guiar el trabajo, la visión de cualquier funcionalidad concreta en las etapas de diseño era la que podría ayudar a mejorar la experiencia de los artistas distribuidos, el trabajo conjunto con los mismos y permitir la participación en las experiencias del diseño y análisis a todos los participantes, incluidos los músicos. También se tuvo que determinar el tipo específico de intérprete —interactor— hacia el cual los diseños deberían ser adaptados. Es así como se propuso obtener una comprensión de varios tipos de músicos, con un enfoque específico en sus interacciones interpersonales y en sus motivaciones.

3. Construcción del test

Después de definir los criterios de evaluación basados en el ensamble, la interacción con los demás, la creatividad y la autoexpresión; y, siguiendo las directrices generales y parámetros impulsores formulados al inicio del proceso de diseño, se encontró que hasta donde sabemos, los métodos estandarizados para evaluar estos factores aún no se han establecido. Herramientas como los cuestionarios de experiencia de juegos abarcan en gran medida preguntas sobre el flujo, la creatividad y la inmersión, esta herramienta demostró ser un elemento adecuado para la evaluación del ensamble y el disfrute debido a la amplitud de los comportamientos examinados (IJsselsteijn, Hoogen, Klimmt, Kort, Lindley, Mathiak, et al., 2008). De hecho, la naturaleza general de las preguntas en los entornos de juego mostraba que era necesaria relativamente poca modificación para adaptarla al contexto musical y audiovisual. Su popularidad y eficiencia motivó el

diseño de los cuestionarios para evaluar puntos de referencia adicionales como la autoexpresión y la creatividad. Para estos cuestionarios también se hizo uso de análisis telemáticos anteriores como los planteados en Telematic Six Perspectives, un paradigmático escrito que propone un cuestionario que, aunque no está estructurado bajo metodologías y técnicas HCI, sí hace uso del amplio conocimiento y experiencia informada de los autores que hacen parte misma de la génesis e implantación de los entornos telemáticos de creación telemática (Oliveros, et al., 2009); de igual forma, fueron usados algunos de los interrogantes y objetivos planteados en los prototipos anteriores, encontrando en ellos importantes variables para ser evaluadas en el prototipo final. Por último se abordaron los parámetros impulsores y las directrices generales de diseño planteadas al inicio del proceso de diseño como guía permanente en el proceso de construcción diseño y evaluación del prototipo.

Es así como se construyó el test para el prototipo MirroR, el cual se encuentra dividido en 4 secciones:

Caracterización del usuario: esta parte del test busca obtener información clara acerca de los participantes a la prueba.

Actuación: este apartado busca conocer cuál es la participación en el *performance* telemático a partir de los criterios descubiertos en el análisis de los prototipos tempranos que permitieron la construcción de roles.

Expectativas: esta sesión busca entender las expectativas de los usuarios a partir de sus perfiles, analizándolos desde sus saberes interpretativos y conocimientos sobre los *performances* telemáticos.

Análisis: los interrogantes planteados en esta sesión buscan dar luces acerca de los procesos y las actuaciones realizadas en la conformación del ensamble y puesta en marcha del *performance* telemático; a continuación, se presentan los ciclos iterativos y pruebas realizadas en cada una de las etapas del *testeo*.

4. Ciclos iterativos y pruebas

Para cada sesión de análisis se contó con dos laboratorios de usabilidad en los cuales se ubican cada uno de los intérpretes, un evaluador de experiencia —*tester*—, un asistente de laboratorio, una videocámara, un sistema de audio estéreo, un redoblante, un monitor de video, un videoprojector, un computador, una interfaz de audio, la interfaz MirroR Box y el prototipo MirroR montado (véase la Figura 0.1 Descripción del montaje del laboratorio de análisis multivariable”).

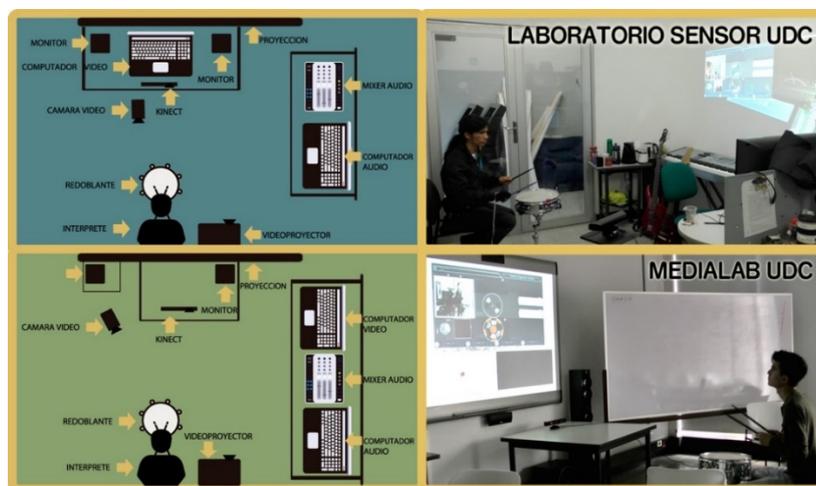


Figura 0.1 Descripción del montaje del laboratorio de análisis multivariable. Fuente: Elaboración propia a partir de la disposición de los laboratorios de usabilidad, junio 2017.

Primer ciclo de testeos

Una vez realizados los ajustes pertinentes se inician los ciclos de testeo. Estos ciclos evolutivos parten de la presentación básica de la interfaz y de la adición sistemática de variables a la interfaz; de esta forma, se inicia el primer ciclo de testeos mutivariables haciendo uso de dos variables, una de las cuales es fija —el sonido en stream—, a partir de esta condición el sistema propone el siguiente orden de testeos a partir del cambio de variables:

Tabla 0.1 Definición del primer ciclo de testeos multivariables. Fuente: Elaboración propia a partir del orden evolutivo de los test planteados, junio, 2017.

Primer Ciclo	Variable 1 - Fija	Variable 2
Test 2	Sonido en <i>stream</i> . JackTrip	Video <i>stream</i> . UG
Test 3	Sonido en <i>stream</i> . JackTrip	Visualización de sensores, giroscopio, acelerómetro y piezoeléctricos. Arduino OSC
Test 4	Sonido en <i>stream</i> . JackTrip	Esqueleto. Kinect OSC
Test 5	Sonido en <i>stream</i> . JackTrip	Visualización 1: nube de puntos
Test 6	Sonido en <i>stream</i> . JackTrip	Visualización 2: superficie de coordenadas

Los datos arrojados al aplicar el primer ciclo de testeos se centran en el uso de las encuestas²⁸ y de las entrevistas semiestructuradas.²⁹ Las encuestas son tabuladas y las entrevistas son analizadas y aplicadas a técnica de conteo de ocurrencias, esta cuantificación de actividades se sintetiza en la tabla siguiente:

Tabla 0.1 Primer ciclo de testeos de cuantificación consolidada. Fuente: Elaboración propia a partir de la cuantificación de categorías analizadas en el primer ciclo de testeos, junio, 2017.

Ciclo	Ensamble	Emoción	Funcionalidad	Interfaz	Interpretación	Sincronía	Colaboración	Performance
1								
Test 2	5,3	5,8	6,5	6,5	6,2	6,0	6,0	6,6
Test 3	6,6	5,8	6,1	6,2	6,5	6,3	6,0	5,2
Test 4	2,5	2,4	4,0	2,5	2,6	2,5	3,2	2,7
Test 5	3,3	5,0	5,5	5,5	4,7	4,5	4,5	3,5
Test 6	3,3	5,2	5,7	5,5	5,4	5,0	4,7	4,8

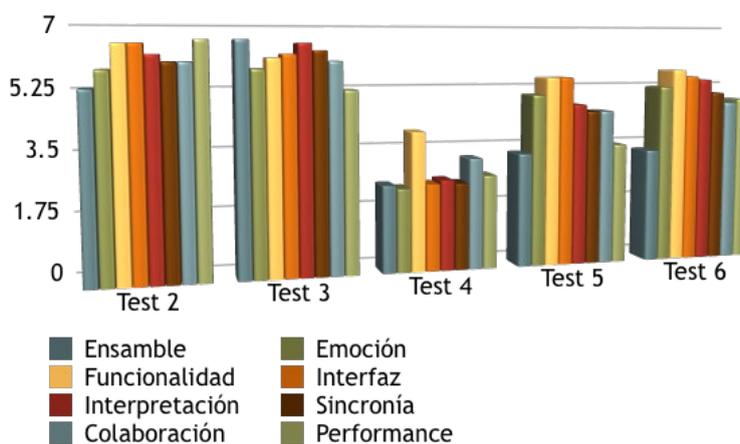


Figura 0.2 Primer ciclo de testeos. Visualización de la cuantificación. Fuente: Elaboración propia a partir de la cuantificación del primer ciclo de testeos, julio 2017.

Después de este primer ciclo de testeo se observaron algunos inconvenientes en el diseño de la interfaz debido a la saturación de objetos en la interfaz, así como la inoperancia de otros objetos; esto hizo replantear, adaptar y rediseñar algunas características y funcionalidades de la interfaz. También se descubrió —gracias a la sugerencia de los intérpretes y el equipo técnico de testeo— que era necesario el diseño de varias interfaces dependiendo del tipo o característica del usuario, es así como se

²⁸ Los datos obtenidos de las encuestas, a las que se hace referencia, pueden ser visualizado en https://drive.google.com/open?id=1DH00I9x_6U5GOamIcJfvtNjraaLNwInB

²⁹ Los videos del primer ciclo de testeos y las entrevistas no estructuradas de este ciclo pueden ser consultados en <http://prototipostesis.blogspot.com.co/2017/03/test-1-del-prototipo.html>

diseñan tres tipos de interfaces y se rediseña la interfaz de intérprete a partir de las sugerencias de las entrevistas semiestructuradas (figura 0.3).



Figura 0.3 Rediseño de interfaz a partir del primer ciclo de testeos. Fuente: Elaboración propia a partir de los cambios realizados en la interfaz MirroR, julio, 2017.

Segundo ciclo de testeos

Después de realizar las mencionadas correcciones en la interfaz, se programó el segundo ciclo de testeos. En este segundo ciclo se trabajó con tres variables que se agruparon, nuevamente, a partir de los objetos de la interfaz. Para la evaluación de este ciclo se hizo uso de las mismas preguntas del primer ciclo y se realizó un segundo ciclo de entrevistas semiestructuradas. Para este ciclo se planteó la combinación exhaustiva de variables (véase la

Tabla 0.3 Análisis exhaustivo multivariable con tres ítems de análisis), pero con miras a no agotar a los intérpretes, al grupo de testeo y a la poca funcionalidad analítica en el uso de variables exhaustivas (Bernsen & Dybkjær, 2009), se seleccionaron para este ciclo las variables presentadas en la

Tabla 0.4 Ciclo dos: test multivariable con tres variables”.

Tabla 0.2 Análisis exhaustivo multivariable con tres ítems de análisis. Fuente: Elaboración propia a partir del orden de testeos realizados con tres variables, julio, 2017.

Segundo ciclo	Variable 1: elementos fijos, sonido + partitura	Variable 2	Variable 3
Test 7	Sonido en <i>stream</i> . Partitura	Video <i>stream</i> . UG	Sensores, giroscopio, acelerómetro y piezoeléctricos. Arduino OSC
Test 8	Sonido en <i>stream</i> . Partitura	Video <i>stream</i> . UG	Esqueleto. Kinect
Test 9	Sonido en <i>stream</i> . Partitura	Video <i>stream</i> . UG	Visualización 1: nube de puntos
Test 10	Sonido en <i>stream</i> . Partitura	Video <i>stream</i> . UG	Visualización 2: superficie de coordenadas
Test 11	Sonido en <i>stream</i> . Partitura	Sensores, giroscopio, acelerómetro y piezoeléctricos. Arduino OSC	Esqueleto. Kinect
Test 12	Sonido en <i>stream</i> . Partitura	Sensores, giroscopio, acelerómetro y piezoeléctricos. Arduino OSC	Visualización 1: nube de puntos
Test 13	Sonido en <i>stream</i> . Partitura	Sensores, giroscopio, acelerómetro y piezoeléctricos. Arduino OSC	Visualización 2: superficie coordenadas
Test 14	Sonido en <i>stream</i> . Partitura	Esqueleto. Kinect	Visualización 1: nube de puntos
Test 15	Sonido en <i>stream</i> . Partitura	Esqueleto. Kinect	Visualización 2: superficie de coordenadas
Test 16	Sonido en <i>stream</i> . Partitura	Visualización 1: nube de puntos	Visualización 2: superficie de coordenadas

Tabla 0.3 Ciclo dos: test multivariable con tres variables. Fuente: Elaboración propia a partir del orden de testeos realizados con tres variables, julio, 2017.

	Variable 1: elementos fijos, sonido + partitura	Variable 2	Variable 3
Test 7	Sonido en <i>stream</i> . Partitura	Video <i>stream</i> . UG	Sensores, giroscopio, acelerómetro y piezoeléctricos. Arduino OSC
Test 8	Sonido en <i>stream</i> . Partitura	Video <i>stream</i> . UG	Esqueleto. Kinect
Test 9	Sonido en <i>stream</i> . Partitura	Video <i>stream</i> . UG	Visualización 1: nube de puntos
Test 11	Sonido en <i>stream</i> . Partitura	Sensores, giroscopio, acelerómetro y piezoeléctricos. Arduino OSC	Esqueleto. Kinect
Test 13	Sonido en <i>stream</i> . Partitura	Sensores, giroscopio, acelerómetro y piezoeléctricos. Arduino OSC	Visualización 2: superficie de coordenadas
Test 16	Sonido en <i>stream</i> . Partitura	Visualización 1: nube de puntos	Visualización 2: superficie de coordenadas

De este nuevo ciclo de test se obtuvo un nuevo registro de datos e información que es presentada en la tabla siguiente.

Tabla 0.4 Segundo ciclo de testeos: cuantificación consolidada. Fuente: Elaboración propia a partir del segundo ciclo de testeos, julio, 2017.

Ciclo 2	Entorno	Emoción	Funcionamiento	Interfaz	Interpretación	Sincronía	Colaboración	Performance
Test 7	6,0	6,8	6,7	6,3	6,3	6,0	5,8	6,0
Test 8	5,3	5,3	5,0	6,0	5,5	5,1	5,5	5,8
Test 9	5,2	5,9	5,5	5,2	5,7	5,4	6,0	6,0
Test 11	5,0	5,4	5,0	5,5	5,7	6,0	5,7	5,8
Test 13	6,6	5,9	6,4	6,3	6,6	6,5	6,2	6,2
Test 16	6,4	5,4	6,2	6,1	6,0	6,0	5,7	6,0

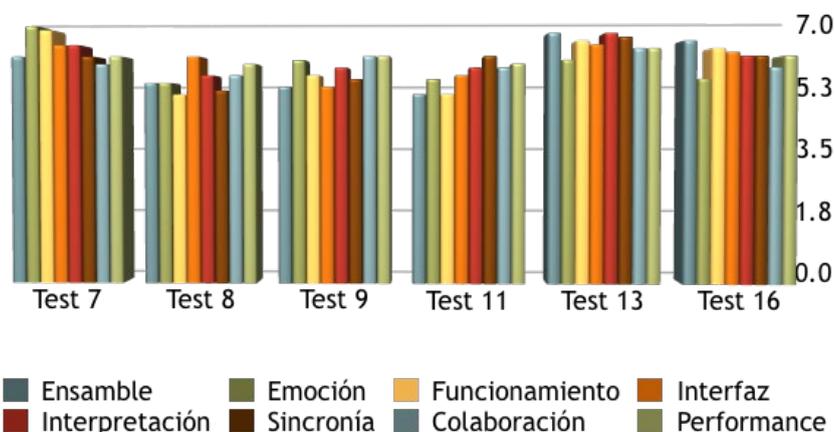


Figura 0.4 Segundo ciclo de testeos: visualización de la cuantificación. Fuente: Elaboración propia a partir de la cuantificación del segundo ciclo de testeos, julio, 2017.

A partir de esta información y con este nuevo ciclo de test se pudo analizar que al aumentar la cantidad de variables en la interfaz, las interacciones de los instrumentistas cambiaba con esta y las relaciones entre los intérpretes también. La relación sonido-video-sensores (test7), junto con la partitura, es en la que más cómodos se sintieron, aunque no fueron los mejores valores de ensamble y performance, como se indica más adelante. En las entrevistas semiestructuradas se evidenció que los cambios realizados en la interfaz fueron altamente beneficiosos y no solo para los intérpretes sino también para el equipo técnico; también facilitó, en gran medida, el diseño del escenario.

Tercer ciclo de testeos

El tercer ciclo de test abordó los siguientes conjuntos de variables.

Tabla 0.5 Test multivariable ciclo tres, con cuatro variables. Fuente: Elaboración propia a partir del conjunto de variables usado en el tercer ciclo de testeos, julio, 2017.

Tercer Ciclo	Variable 1 - elementos fijos + Sonido + partitura	Variable 2	Variable 3	Variable 4
Test 17	Sonido en stream. Partitura	Video stream. UG	Sensores, giroscopio, acelerómetro piezoeléctricos. Arduino OSC	Visualización 2: nube de superficie de coordenadas
Test 18	Sonido en stream. Partitura	Video stream. UG	Esqueleto. Kinect	Visualización 1: nube de puntos
Test 19	Sonido en stream. Partitura	Sensores, giroscopio, acelerómetro y	Visualización 1: nube de puntos	Visualización 2: nube de superficie de coordenadas

		piezoeléctricos. Arduino OSC		
Test 20	Sonido en <i>stream</i> . Partitura	Sensores, giroscopio, acelerómetro y piezoeléctricos. Arduino OSC	Esqueleto. Kinect	Visualización 2: nube de superficie de coordenadas

En este nuevo ciclo se simplificaron las preguntas del test en razón a que afloraron algunas categorías de mayor interés en el transcurso de las pruebas anteriores; adicionalmente, algunas preguntas fueron muy repetitivas en el transcurso de los laboratorios ya que las respuestas se consolidaban una y otra vez entre los diferentes test. Por lo anterior, el enfoque de las preguntas se centró en aquellas que brindaran nueva información entre los test y permitieran analizar de forma sintética los cambios observados por los participantes en la evolución de la interfaz. La selección no exhaustiva de las variables también se simplificó a partir de las relaciones entre variables con mejores expectativas, seleccionadas estas a partir de nuevas preguntas, surgidas en el diálogo con los participantes en el montaje de cada test, es de anotar que estos aportes dados por los participantes en los laboratorios, así como los cambios sobre la marcha, nos brindaron gran conocimiento, no solo sobre el prototipo a evaluar sino sobre el método mismo de evaluación impulsada por el usuario.

Fruto de este nuevo ciclo de test y entrevistas semiestructuradas se obtuvieron los datos de la tabla siguiente.

Tabla 0.6 Tercer ciclo de testeos cuantificación consolidada. Fuente: Elaboración propia a partir de la cuantificación de los datos obtenidos en el tercer ciclo de testeos, julio, 2017.

Ciclo	Ensamble	Emoción	Funcionamiento	Interfaz	Interpretación	Sincronía	Colaboración	Performance
Test 17	6,8	6,5	6,8	6,7	6,5	6,5	6,5	6,8
Test 18	5,4	5,0	5,4	5,1	5,2	5,5	5,1	5,1
Test 19	6,6	6,5	6,7	6,5	6,2	6,5	6,0	6,1
Test 20	5,6	2	5,4	5,3	5,8	5,5	5,3	5,1

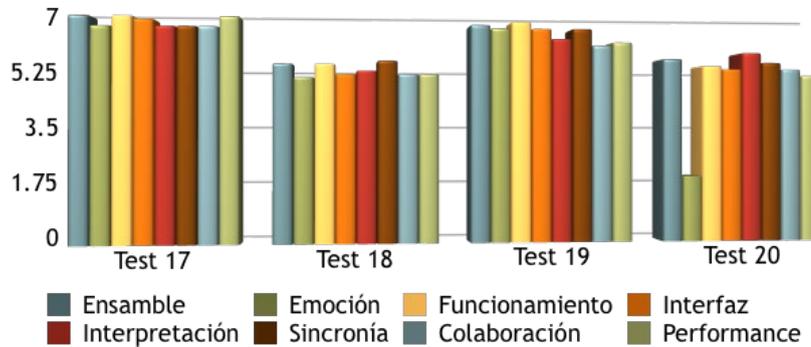


Figura 0.5 Tercer ciclo de testeos: visualización de la cuantificación. Fuente: Elaboración propia a partir de la cuantificación de los datos obtenidos en el tercer ciclo de testeos, julio, 2017.

No sin sorpresa encontramos que el test que arrojó mejores características tanto en la interpretación, como en el ensamble y el performance fue el test 17, el cual sintetiza de alguna forma las mejores variables descubiertas en los ciclos anteriores, es de resaltar algunas de las respuestas brindadas por los participantes donde se observa que cada vez más prestan menos atención al stream de video y se concentran más en la representación de los sensores y en las visualizaciones, encontrando incluso que inflexiones específicas de la interfaz funcionaban como diálogos concretos en la interacción entre los instrumentistas, tal vez por esto también encontramos la cercanía en la evaluación entre el test17 y el test19.

Cuarto ciclo de testeo

El cuarto y último ciclo de test conjugó el uso de 5 y 6 variables, las variables evaluadas en los test de este ciclo se presentan a continuación.

Tabla 0.7 Test multivariable: ciclo 4 uso de 5 y 6 variables. Fuente: Elaboración propia a partir de la selección de variables para el cuarto ciclo de testeos, julio, 2017.

Cuarto Ciclo	Variable 1: elementos fijos sonido + partitura	Variable 2	Variable 3	Variable 4	Variable 5	Variable 6
Test 21	Sonido en <i>stream</i> . Partitura	Video <i>stream</i> . UG	Sensores, giroscopio, acelerómetro y piezoeléctricos. Arduino OSC	Visualización 1: nube de puntos	Visualización 2: superficie de coordenadas	-
Test 22	Sonido en <i>stream</i> . Partitura	Video <i>stream</i> . UG	Sensores, giroscopio, acelerómetro y piezoeléctricos. Arduino OSC	Visualización 1: nube de puntos	Esqueleto. Kinect	-
Test 23	Sonido en <i>stream</i> . Partitura	Video <i>stream</i> . UG	Visualización 1: nube de puntos	Visualización 2: superficie de coordenadas	Esqueleto. Kinect	-
Test 24	Sonido en <i>stream</i> . Partitura	Video <i>stream</i> . UG	Sensores, giroscopio, acelerómetro y piezoeléctricos. Arduino OSC	Visualización 1: nube de puntos	Visualización 2: superficie de coordenadas	Esqueleto. Kinect

En este último ciclo de test se consolidaron algunas de las apreciaciones y hallazgos realizados en los ciclos anteriores, este ciclo fue de gran utilidad ya que en este los participantes dieron apreciaciones generales y propusieron opciones y caminos para desarrollos futuros, la mayor utilidad de este ciclo se dio en los análisis relacionados con los montajes y el performance para el público. Si bien estas preguntas estuvieron siempre en los cuestionarios de los test, fue en este último ciclo donde tanto las preguntas como las entrevistas brindaron más información sobre este campo, es así como las evaluaciones sobre los últimos test reflejan que el ensamble y el montaje fue mejor para el test 24, mientras que tanto el test 24 como el 21 mostraron ser iguales para la sincronía ya que el esqueleto no da valor agregado en esta categoría. En este ciclo se demostró por parte de los intérpretes que el entendimiento de los objetivos de la interfaz y el reconocimiento de la misma permite fomentar en gran medida el ensamble del grupo y que gracias a la presencia de la interfaz como un elemento más en el proceso de construcción y diálogo telemático generó diferentes tipos de actividades y disposiciones que no hubieran existido con un simple modelo de videoconferencia, los participantes también fueron claros en manifestar y evaluar que este cambio en la construcción del performance motivó su participación y frente a la obra, permitiéndoles comprender el entorno telemático como un campo nuevo y diferente que debe ser abordado desde perspectivas diferentes a las tradicionales. Los intérpretes también plantearon la posibilidad de avanzar más en la representación del gesto y en los modelos predictivos de interpretación que de alguna forma admitan

trabajar de manera diferente con la latencia, más allá del problema de sincronía y temporalidad.

Tabla 0.8 Cuarto ciclo de testeos: cuantificación consolidada. Fuente: Elaboración propia a partir de la cuantificación de los datos obtenidos en el cuarto ciclo de testeos, julio, 2017.

Ciclo 4	Ensamble	Emoción	Funcionamiento	Interfaz	Interpretación	Sincronía	Colaboración	Performance
Test 21	6,7	6,9	6,9	6,9	6,7	6,5	6,8	6,8
Test 22	6,2	6	6,3	5,7	5,9	6,1	5,2	6,3
Test 23	6,3	6,3	6,3	5,7	6,2	6	5,5	6,5
Test 24	6,9	6,9	6,8	6,9	6,7	6,5	6,8	6,8

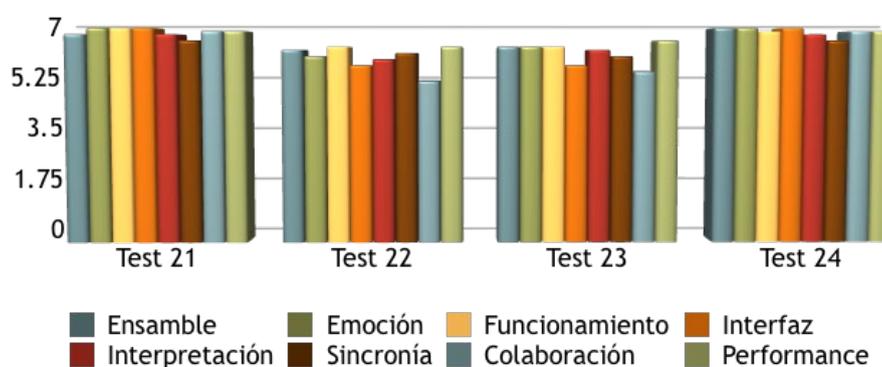


Figura 0.6 Cuarto ciclo de testeos: visualización cuantificación. Fuente: Elaboración propia a partir de la cuantificación de los datos obtenidos en el cuarto ciclo de testeos, julio, 2017.

Categorización de información, destilando la información

Con miras a aumentar la colaboración entre los intérpretes desplazados y con el objetivo de analizar la construcción y apropiación de los tele-espacios activos, fueron diseñados estos entornos de respuesta y las diferentes estrategias de análisis y evaluación para el performance distribuido. Es así como, mediante el uso de cuatro directrices o características planteadas, a saber: el performance, el ensamble, la interacción y la colaboración, y el tamizado, a partir de los elementos que propusimos para la construcción del espacio activo: interfaz, visualización y emoción, analizamos cómo los intérpretes pueden trabajar, crear y alterar, de forma simple el contexto del performance telemático mismo, al responder interactiva y colectivamente. Este sistema permitió a los intérpretes utilizar la interface y todas las funcionalidades del sistema sin tener que separarse de la tarea principal: su interpretación colectiva. Además, permitió a la investigación analizar la labor de los intérpretes en la construcción del tele-espacio activo, si bien las conclusiones de cada etapa de interrelación entre los intérpretes está sintetizada en los apartados anteriores que se han ido destilando conforme ha avanzado

este capítulo, podríamos resumir estos hallazgos y muestras, en la siguiente tabla.

Tabla 0.9- Hallazgos finales, análisis de la información. Fuente: Elaboración propia a partir de los hallazgos encontrados en el proceso, julio, 2017.

Directrices	Elementos	Previo	Alcanzado
<i>Performance</i>	Interfaz	Metáforas de videoconferencia	Interfaces de representación gestual y predictiva
	Visualización	Stream de video. Masa de movimiento	Comprensión y diseño del escenario activo
	Emoción	Satisfacción al interpretar música compartida	Deseo de participar en nuevos <i>performances</i> telemáticos
	Interpretación	Lectura de partitura	
Ensamble	Interfaz	Reconocimiento de la interfaz	Inflexiones específicas de la interfaz
	Visualización	Síntesis gráfica	Reconocimiento de gestos e intenciones a partir de visualizaciones
	Emoción	Análisis y sincronía y retardos	Disfrute de la interpretación y el ensamble
	Interpretación	Partitura	Apropiación y lectura de la interfaz
Interacción	Interfaz	Objetivos de la interfaz	Modos de control individual y grupal
	Visualización	Visualización de los sensores	La forma, en las visuales, estaba directamente relacionada con la posición del brazo. Percepción visual del gesto
	Emoción	Frustración en control de la interfaz	Conciencia del gesto y su representación en la interfaz
	Interpretación	Lectura de partitura, control del tiempo y sincronía	Interfaz directamente proporcional a las acciones de los intérpretes
Colaboración/colectivo	Interfaz	Lectura de la interfaz	Interpretación
	Visualización	Atención en los detalles	Construcción conjunta, trabajo co-lectivo
	Emoción	Baja colaboración	Actitud participativa y propositiva
	Interpretación	Acuerdos previos	Señales interpretativas y expresivas

Con una serie de técnicas impulsadas por el usuario, nuestro entorno de respuesta evolucionó a través de varias etapas para llegar a su representación actual. Siguiendo las pautas del diseño centrado en el usuario, se comenzó con un enfoque en el usuario mediante observaciones, entrevistas y perfiles de persona que, a su vez, se utilizaron para obtener comentarios de los usuarios mediante testeos formales, en los cuales, el nivel de participación de los usuarios fue en aumento; primero, con la evaluación inicial del prototipo y luego con la colaboración de los laboratorios de análisis, incorporando

sus comentarios y sugerencias en cada etapa. Cuestionarios recogidos a lo largo de la colaboración demostraron que nuestro sistema ayudó a mejorar el sentido de los intérpretes en cuanto al disfrute y autoexpresión. Además, el análisis cualitativo de nuestras discusiones con los miembros de los laboratorios, ha demostrado que encontraron el sistema práctico, y que probablemente lo usarían de nuevo.

Conclusiones

En términos generales, se logró el desarrollo de un entorno performático audiovisual telemático, un tele-espacio activo, donde los intérpretes pueden interactuar telemáticamente participando activamente en la construcción del ensamble y experimentar un nuevo campo performático aumentado.

Dentro del desarrollo de los diferentes performances telemáticos se fue categorizando el entorno virtual generado por la confluencia no solo de los espacios remotos sino de las interacciones generadas a partir del que hacer de los intérpretes, los creadores y los técnicos, de esta forma el concepto de espacio en el espacio activo refiere un lugar virtual que es designado o define un nuevo tipo de espacio simbolizado, dicho espacio se conforma teniendo en cuenta lo que se percibe y se desarrolla en el gracias a la relación comunicativa generada por la interacción entre los participantes, es así como esta información perceptual del espacio se construye a partir de la conformación de la interfaz y de la interacción sobre la misma, en el caso del performance se da gracias a la gestualidad del interprete virtualizado a partir del uso de sensores y de su representación en la interfaz embebida (o espacio mediatizado), generándose así un concepto de presencialidad que se desarrolla a partir de la interacción gestual y simbólica sobre la interfaz y en últimas sobre la construcción misma del espacio activo. Esta mirada de la conformación del espacio telemático y de la potencia del gesto como presencia en la construcción de la interfaz es una de las conclusiones laterales más importantes planteadas en esta investigación.

Entender el lugar como el espacio del ensamble, donde los participantes tejen una estructura comunicativa y sobretodo social y donde se establecen tanto lazos como vínculos interpretativos, performáticos y afectivos, brinda al tele-espacio activo una característica de entorno más allá de lo mediatizado y mucho más complejo e interesante que el simple hecho de la telecomunicación, el compartir esta construcción de lugar es uno de los puntos de vista que aporta esta investigación frente a la construcción del performance telemático musical.

El uso de una pieza de creación musical como un experimento que valida las hipótesis de la investigación dentro de un proceso estructurado como un entorno empírico analítico, es una propuesta que, si bien hace parte del concepto investigación creación, se aleja de la estructura comúnmente dispuesta en la que el producto de investigación es una obra de creación, al hacer uso de la creación de obras dentro del proceso de investigación y no únicamente como un producto final se plantea, a partir de esta investigación, una interesante alternativa metódica en el campo de la investigación creación, que puede ser utilizada en el desarrollo de diferentes investigaciones del Doctorado.

Agradecimientos

Este trabajo ha sido financiado parcialmente por el Proyecto de Investigación "Espacios Telemáticos" y forma parte de la Tesis doctoral "Configuración de tele-espacios activos. Entornos de interacción performáticos distribuidos de creación sonora y visual".

Referencias

- 1 Bernsen, N., & Dybkjær, L. (2009). *Multimodal Usability*. Verlag, London, UK: Springer.
- 2 Doering, T., Pfleging, B., Kray, C., & Schmidt, A. (2010). Design by Physical Composition for Complex Tangible User Interfaces. *Proceeding CHI EA '10 CHI '10 Extended Abstracts on Human Factors in Computing Systems* (pp. 3541-3546). New York: ACM Publishher.
- 3 Jordà, S. (2002). FMOL: Toward User-Friendly, Sophisticated New Musical Instruments. *Computer Music Journal*, 26(3), 23-39.
- 4 MacDonald, C., & Atwood, M. (2013). Changing perspectives on evaluation in HCI: past, present, and future. *CHI EA '13 CHI '13 Extended Abstracts on Human Factors in Computing Systems* (pp. 1969-1978). New York: Publisher ACM.
- 5 Padhye, N., & Kalia, S. (2009). Rapid prototyping using evolutionary approaches. *Proceeding GECCO '09 Proceedings of the 11th Annual Conference Companion on Genetic and Evolutionary Computation Conference: Late Breaking Papers* (pp. 2725-2728). New York: ACM Publisher.
- 6 Wood, L. (1997, March). Semi-structured interviewing for user-centered design. *Magazine interactions*, 4(2), 48-61.

Centro de Análisis de Datos y Supercómputo (CADS) de la Universidad de Guadalajara

Luis Alberto Gutiérrez Díaz de León, Jorge Lozoya Arandia, Lizette Robles Dueñas y Claudia Patricia Franco Sánchez

Coordinación General de Tecnologías de Información, Universidad de Guadalajara,
Av. Juárez 976, Piso 2, Col. Centro, C.P. 44100, Guadalajara, Jalisco, México.

luis.gutierrez@redudg.udg.mx, jorge.lozoya@redudg.udg.mx, lizette@redudg.udg.mx,
claudia.franco@redudg.udg.mx

Resumen: El presente documento describe el proyecto de implementación del Centro de Análisis de Datos y Supercómputo (CADS) en la Universidad de Guadalajara, el cual es un espacio que se creó para apoyar el desarrollo de la investigación entre grupos multidisciplinarios en donde interactúa parte de la comunidad científica, tecnológica y académica y la generación de nuevo talento que demanda el uso del supercómputo y del análisis de datos, contribuyendo a posicionar al país como referente de innovación. El CADS apoya los procesos de innovación y el desarrollo de investigaciones científicas de calidad para la comunidad universitaria, gracias al procesamiento y al análisis de grandes volúmenes de información en sistemas de cómputo de alto rendimiento, en un ámbito multidisciplinario, interdisciplinario y transdisciplinario; a su vez, cuenta con la participación de personal especializado en el mantenimiento y operación de la infraestructura tecnológica, además de contar con un nivel de pericia técnica que le permite ser el nexo especializado entre el equipo de alto rendimiento y el investigador. La implementación de este espacio conllevó una serie de esfuerzos tanto tecnológicos como de implementación de procesos, para entregar el servicio de procesamiento y análisis de datos, apegado a las mejores prácticas de implementación de servicios de TI. La gestión y visión de la alta dirección fue uno de los factores importantes para la implementación, siendo identificada la necesidad de contar con estas herramientas tecnológicas ante la realidad que actualmente enfrentan los procesos de generación de conocimiento, y debido a la gran cantidad de datos que se generan segundo a segundo en todas las ramas de la ciencia, y aunado a una serie de planes y procesos para su implementación, lo que permitió llevar a la comunidad universitaria y a la región occidente de México a habilitar este espacio tan importante para el desarrollo de la ciencia.

Palabras clave: supercómputo, análisis de datos, Big Data, procesamiento.

Eje temático: Tecnologías en la innovación educativa e investigación.

1. Introducción

La Universidad de Guadalajara es la segunda institución de educación superior más grande en México con una matrícula de más de 287 mil 760 estudiantes distribuidos en el estado de Jalisco. Cuenta con 15 centros universitarios, 6 de ellos son temáticos (disciplinas enfocadas a una rama) y 9 se encuentran en zonas regionales siendo multidisciplinarios. Cuenta con una plantilla de 16,854 académicos. Es la universidad con 574 cuerpos académicos consolidados en el Programa para el Desarrollo Profesional Docente (PRODEP), y es la segunda universidad del país con más profesores miembros del Sistema Nacional de Investigadores (SNI), con una numeraria de 1,293 registrados. La Universidad de Guadalajara cuenta con más de 184 posgrados inscritos en el Padrón Nacional de Posgrados de Calidad (PNPC), lo que la posiciona como la segunda universidad a nivel nacional con mayor número de posgrados inscritos a este programa.

Algunos de los programas educativos de posgrado con los que cuenta la Universidad de Guadalajara demandan el uso y desarrollo de sistemas de cómputo de alto rendimiento, como una herramienta de apoyo en las investigaciones, principalmente de académicos adscritos en aquellos centros universitarios afines a las líneas de análisis antes mencionadas, como el de Ciencias Exactas e Ingenierías (CUCEI), el de Ciencias Biológicas y Agropecuarias (CUCBA), el de Ciencias Sociales y Humanidades (CUCSH), el de Ciencias Económico Administrativas (CUCEA), el de Ciencias de la Salud (CUCS), el Centro Universitario de Tonalá (CUTonalá) y el Centro Universitario de los Lagos (CULagos), pero sin descartar aquellos centros en donde sus investigadores se puedan ver beneficiados al considerar y explotar las potencialidades del supercómputo en sus propios trabajos de investigación.

Ante esta necesidad la Universidad de Guadalajara, a través de la Coordinación General de Tecnologías de Información (CGTI) hizo un gran esfuerzo por consolidar un espacio de supercómputo con un equipo de alta calidad y velocidad con las tecnologías más avanzadas disponibles en el mercado, para que con ello impere y se extienda en la Universidad de Guadalajara la cultura de uso del supercómputo en las investigaciones que por su nivel de complejidad así lo requieran, y que puedan competir con otras investigaciones en calidad, precisión e innovación.

Dentro de este contexto como comenta Schwab "la actividad investigativa y la innovación en tecnología es de gran relevancia para el desarrollo económico de las sociedades; por ello, es de suma importancia que, desde las instituciones de educación superior, laboratorios de estudios avanzados o específicos, y centros de investigación, se formalice y apoye esta actividad entre los académicos y estudiantes.^[1] De esta manera entregar un servicio de supercómputo y análisis de datos como apoyo en el desarrollo de investigaciones de alto nivel se identificó como una prioridad institucional.

2. Contexto y problemática en que se propuso el proyecto

Esfuerzos de supercómputo en la Universidad de Guadalajara

La Universidad de Guadalajara a lo largo de su historia ha sido una institución que se ha preocupado por ir al ritmo de los cambios sociales y modelos educativos, y es de la misma manera en el uso de tecnologías de información; estos cambios se basan en el Plan de Desarrollo Institucional (PDI) 2014-2030 de la Universidad de Guadalajara. Una de las líneas de trabajo que se han desarrollado es contar con servicios de supercómputo, años atrás se trazaron esfuerzos para brindar servicios de supercómputo, entre ellos, se realizaron inversiones para la adquisición de equipos y se creó una estructura para ofertar este importante servicio. Desde 1994, la entonces Coordinación General de Sistemas de Información (CGSI), contaba con el Centro de Cómputo de Alto Rendimiento (CENCAR), a través del cual tenía las funciones de constituir y administrar dicho centro, impulsar acciones de vinculación de dicho centro con los órganos y entidades de la Red Universitaria, así como ofrecer servicios de cómputo de alto rendimiento a la comunidad universitaria, tal como lo establecía la fracción X del artículo 11 del Reglamento Interno de la Administración General de la Universidad de Guadalajara.

Sin embargo, cuando se extinguió la CGSI en 2007, la mayor parte de sus atribuciones fueron conferidas a la actual Coordinación General de Tecnologías de Información (CGTI), sin que las atribuciones del CENCAR se hicieran explícitas en la estructura de la CGTI.

En 2007 fue creada la Coordinación General de Tecnologías de Información (CGTI), la cual ha contado con una infraestructura compuesta por equipos de cómputo de procesamiento intensivo, dispositivos de almacenamiento masivo y redes de alta velocidad, como herramientas de apoyo para la investigación científica y tecnológica. Derivado de ello, el CADS de la Universidad de Guadalajara se crea como un espacio de cálculo avanzado que requiere de un equipo sofisticado de operación y análisis de datos, así como de recurso humano capacitado para el manejo y mantenimiento del mismo; en el año 2018 se implementa como un espacio que formaría parte de la propia estructura organizacional y operativa de la CGTI.

Problemática

La iniciativa de creación de un centro de análisis de datos y supercómputo comenzó hace 5 años; esta inquietud nace de la actualidad tan acelerada que vivimos en cuanto a lo tecnológico, el avance en las tecnologías de cómputo y análisis de datos en las últimas tres décadas, lo cual ha permitido la generación de información a un ritmo tan acelerado como no se había presentado en la historia de la humanidad. Los avances de las redes de datos nos han permitido compartir esta información alrededor del mundo prácticamente al instante en el que se generan los datos, esto ha transformado la vida diaria para toda la sociedad y se ha reflejado en los procesos de investigación y avances en la ciencia. Contar con capacidades de procesamiento cada vez mayores permite

obtener resultados en tiempos tan cortos que hace algunos años eran inimaginables; estos resultados han sido aplicados con la misma inmediatez en todo tipo de procesos productivos, políticos, sociales y otros más.

"El supercómputo en la actualidad es toral para el desarrollo de la ciencia y tecnología. La supercomputadora es, hasta ahora, la tecnología más vanguardista y robusta de cálculo en el mundo, pues es capaz de resolver, en menor tiempo, desafíos técnicos y científicos (puede atender millones de instrucciones por segundo), respecto de una computadora convencional. Es un sistema que está conformado por componentes de procesamiento trabajando en conjunto para incrementar su capacidad de almacenamiento, requerida para las labores de investigación profunda de las distintas universidades e industrias del mundo". [2]

La implementación de un espacio con estas tecnologías implica un esfuerzo tecnológico que involucra ingenierías de centros de datos, redes de alta capacidad, procesamiento, almacenamiento e implementación de software. Además del gran reto presupuestal que permita esta implementación y de gestión ante las diferentes instancias directivas para determinar la pertinencia y viabilidad del proyecto.

"Los principales retos para habilitar a las universidades de un equipo tan sofisticado como es el supercómputo, son las importantes inversiones que deben realizarse –gestión de políticas públicas– para su adquisición, habilitación, infraestructura, uso y mantenimiento, pero también para el fortalecimiento de la conectividad de redes avanzadas". [3]

La formación de personal altamente calificado que pueda operar y aprovechar un equipo de supercómputo, así como de recursos humanos que sean el vínculo especializado y multidisciplinario entre las funcionalidades del equipo y las propias necesidades de los investigadores, con la finalidad de aprovechar las posibilidades reales de análisis del equipo, se convierte en una de las prioridades para hacer que el proyecto logre sus metas.

A la par de estas implementaciones y capacitaciones se elaboraron una serie de lineamientos y políticas de operación que permitieron generar procesos y procedimientos apegados a estándares de buenas prácticas en el uso y aprovechamiento de tecnologías de información. Esto permite llevar de manera adecuada los proyectos que se procesan y analizan en el centro, al generar métricas que permiten evaluar el aprovechamiento del equipamiento y la pertinencia de los trabajos a procesar en el CADS.

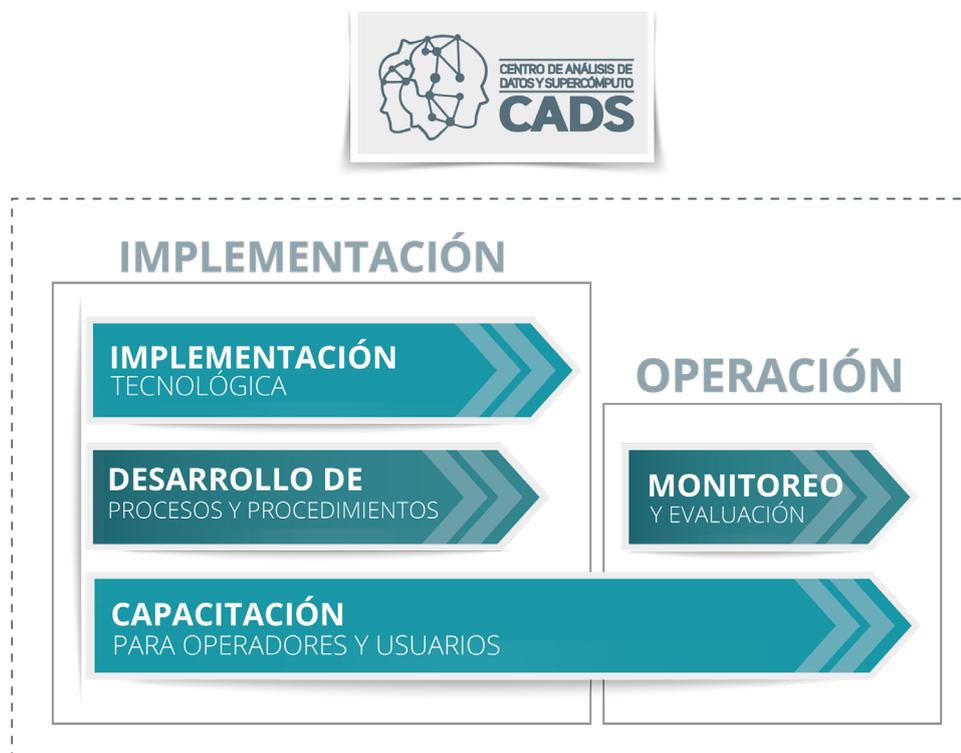


Fig. 1. Proceso de implementación del CADS.

De esta manera la Coordinación General de Tecnologías de Información de esta casa de estudio, generó un esfuerzo de gran magnitud que con estas premisas permitió implementar un espacio con las herramientas de supercómputo y análisis de datos; mismas herramientas que a nivel mundial son cada vez más fundamentales para el desarrollo científico.

3. Descripción de la solución tecnológica implementada

La CGTI elaboró un plan de implementación para la habilitación tecnológica del equipo de supercómputo, y a la par se elaboró una serie de procedimientos que tiene como fin la entrega del servicio de procesamiento y análisis de datos. Estas actividades contemplaron una serie de capacitaciones en el manejo de las tecnologías que se implementaron para el personal que está a cargo de la operación del servicio.

Habilitación del equipo de supercómputo

En el CADS se alberga la supercomputadora más poderosa del país llamada Leo Átrox ("el león de gran cerebro", de acuerdo con el significado de su nombre científico). Este equipamiento requiere una serie de tecnologías de centros de datos para mantener su disponibilidad, incluyendo redundancias en soporte electrónico de datos y de aires acondicionados, así como una serie de equipos que mantenga seguras las instalaciones. Es un espacio con tecnologías de gran rendimiento asociadas con el cálculo, el almacenamiento y la comunicación, cuyo propósito es posibilitar y acelerar las investigaciones científicas y el desarrollo tecnológico de la comunidad universitaria, considerando para ello los diferentes trabajos de investigación y la optimización de los procesos de los mismos.

Las características de Leo Átrox son las siguientes:

- 150 nodos cómputo con un rendimiento mayor a **500 Tflops** y más de **5,000 cores**
- 2 sistemas de almacenamiento con una capacidad mayor a **1.5 PB**
- 1 sistemas de red de baja latencia OPA (Omni-Path) a una velocidad de **100 Gbps**

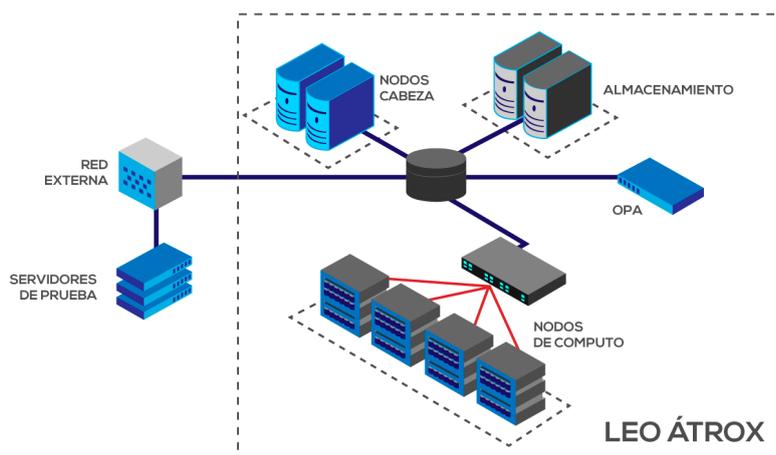


Fig. 2. Arquitectura de Leo Átrox.

El poder de cómputo de Leo Átrox equivale a tener 300,000 celulares de cuarta generación en uso, es decir, todos los estudiantes de la Universidad de Guadalajara realizando al mismo tiempo una operación matemática.

El equipo de supercómputo está soportado en un centro de datos que cuenta con un sistema de respaldo eléctrico con redundancias, por medio de una planta generadora de energía de emergencia y un sistema de baterías capaz de mantener la disponibilidad del servicio sin depender del suministro eléctrico por parte de la Comisión Federal de Electricidad (CFE). Este centro de datos cuenta también con sistemas de aire acondicionado capaz de soportar con doble redundancia la temperatura y humedad dentro del mismo; además de un sistema especializado antiincendios. Es importante mencionar que debido a que la ubicación del edificio del CADS está situada en una

zona sísmicamente activa, se instaló en cada gabinete que aloja el equipo de supercómputo un sistema que protege el equipo contra vibraciones sísmicas.

Desarrollo del proceso de entrega de recursos de cómputo y análisis de datos

La Coordinación General de Tecnologías de Información (CGTI) ha definido políticas para ofertar el servicio del equipo de supercómputo del CADS de la Universidad de Guadalajara, las cuales tienen como objetivo regular los procedimientos aplicados en las solicitudes de registro, asignación de recursos, de acceso, manejo y almacenamiento de datos, soporte técnico y capacitación.

EL CADS inicia operaciones con un catálogo de servicios que dio pie a el desarrollo de los procesos y procedimientos que describen el flujo de actividades y responsables de su atención, todo acorde con los lineamientos del sistema de gestión de seguridad y gobierno de la CGTI.

Servicios ofertados en el CADS:

- Procesamiento de datos e imágenes
- Configuración de aplicaciones científicas
- Desarrollo y optimización de código
- Modelos de paralelización de software
- Soluciones de análisis de grandes volúmenes de datos (Big Data)
- Visualización científica
- Servicios de consultoría para elaboración de proyectos de investigación
- Capacitación

Cada proceso asociado con los servicios ofertados se realizó mediante una serie de análisis y trabajos en conjunto con los responsables de la ejecución del servicio y la alta dirección, esto con el fin de identificar de primera mano tanto las necesidades para entregar el servicio, como darle el enfoque y visión con el que opera el CADS.



Fig. 3. Diagrama del proceso de atención de solicitud de procesamiento.

Plan de capacitación

La capacitación del personal y académicos que hacen uso de los servicios del CADS, así como de toda la comunidad universitaria es uno de los principales ejes su misión como centro. De esta manera se desarrolló un plan de capacitación que contempla los diferentes perfiles a quienes están dirigidos los servicios y conocimientos que se desarrollan en el centro. Los perfiles identificados a quien se dirige la capacitación son usuarios principiantes, usuarios especializados y operadores de equipo de supercómputo; este último se divide en administradores de aplicaciones y administradores de hardware.



Fig. 4. Capacitación continua del CADS.

Las diferentes líneas de capacitación son dinámicas y responden a las necesidades de implementación y desarrollo de las diferentes aplicaciones, sistemas operativos y algoritmos como apoyo a la investigación.

El principal objetivo del CADS es contribuir con el crecimiento y desarrollo de la ciencia, por medio de tecnologías de alto rango de todo México, aunque es dependiente de la Universidad de Guadalajara, las puertas están abiertas para cualquier institución tanto pública como privada.

4. Resultados obtenidos y su impacto

El Centro de Análisis de Datos y Supercómputo (CADS), fue inaugurado el 15 de octubre del 2018. Es un espacio en donde converge la innovación para el desarrollo y procesamiento de datos con la más actual tecnología de supercómputo; en este espacio podemos encontrar personal altamente calificado en análisis de datos, mismo que apoya directamente al desarrollo de investigaciones de alto nivel que requieren procesamiento de gran cantidad de datos y su análisis, transformando con las tecnologías de información y comunicación las prácticas de investigación y educativas para abonar a los distintos campos de la ciencia.

Leo Átrox inició sus operaciones con ocho investigaciones en el área médica, biológica, energética y astronómica, generando proyectos de tecnologías emergentes con aplicación dentro de la Universidad de Guadalajara, como blockchain, inteligencia artificial y analíticos. Brinda a los investigadores, así como a la industria, diversos servicios como el procesamiento de datos e imágenes, configuración de aplicaciones de cálculo numérico, aplicaciones científicas, optimización de código orientado a la paralelización, acompañamiento para análisis de grandes volúmenes de datos (Big Data), visualizaciones científicas, *rendering*, capacitación especializada por parte de expertos, entre otras actividades.

En esta primera etapa de operación se han realizado diversas actividades de difusión y capacitación, y diferentes investigaciones han sido apoyadas en diferente medida en el CADS, por ejemplo:

- Estudio sistemático de pequeñas moléculas antioxidantes hasta estructuras más grandes de carbono.
- Investigación en el área de la astrofísica y física de astropartículas y a la vez en la astronomía multimensajero.
- Centro de datos verdes utilizando energía renovable.
- Identificación del uso de la red inalámbrica institucional
- Investigación enfocada en la eficiencia académica, a través de un análisis estadístico descriptivo y análisis estadístico predictivo para entender el comportamiento del alumnado en su estancia universitaria.
- Investigación en el área de modelado de materiales termoeléctricos.
- Investigación enfocada en las propiedades magnéticas de los cúmulos de átomos.
- Línea de investigación de la geografía económica, donde se busca la medición del crecimiento económico desde el espacio exterior.
- Análisis del comportamiento impulso-respuesta del mercado accionario norteamericano.
- Análisis del modelo estocástico del equilibrio general
- Línea de investigación enfocada en la relatividad general y la astrofísica
- Detección de hallazgos para imágenes para análisis médicos mediante inteligencia artificial.
- Investigación genómica del cáncer y toxicogenómica, línea de generación de conocimiento: biomedicina, marcadores moleculares y salud ambiental en enfermedades poligénica y multifactoriales.
- Análisis de imágenes histopatológicas.

- Análisis de redes genéticas en cáncer

Estas líneas de investigación son apoyadas en diferentes aspectos, desde la adecuación del código implementación de los ambientes de pruebas para sus desarrollos y el despliegue en los nodos de cómputo para su procesamiento, hasta su análisis una vez obtenidos los resultados, su análisis.

5. Aprendizajes (aciertos y errores)

La implementación de un sistema de supercómputo implica un reto complejo tanto tecnológico como de capacidades y conocimiento de personal. La puesta en operación del CADS en la Universidad de Guadalajara, junto con la habilitación del equipo de supercómputo representó un gran reto de coordinación entre las etapas y personas involucradas de diferentes áreas, desde el proceso de gestión administrativa, la habilitación de un espacio físico, así como la puesta en operación del equipo; aunque uno de los retos más importantes es el poder contar con el personal altamente calificado que pueda atender la demanda de los usuarios que requieren este tipo de tecnología.

El trabajo realizado previo a contar con el equipo funcionando, implicó un esfuerzo significativo, el cual podríamos clasificarlo en 4 etapas donde cada una de ellas representan un aprendizaje:

- Identificación de las necesidades de la comunidad para la selección del equipo de supercómputo
- Habilitación del centro de datos y la supercomputadora
- Implementación de los procesos de operación
- Capacitación del personal

La Universidad de Guadalajara cuenta con una gran cantidad de investigadores y alumnos de posgrado en las diferentes áreas de la ciencias como física, geografía, química, genética, bioinformática, por mencionar algunas, además de carreras relacionadas con la ciencias computacionales y otras áreas que demandan el uso de infraestructura de supercómputo en la realización de los proyectos, por lo que significó todo un reto poder identificar las necesidades específicas de cada uno de ellos: uno por la cantidad de personas que se tendría que tomar en cuenta y otra por el reto geográfico que esto representa; por lo que después de realizar un levantamiento al interior de toda la Universidad, se invitó a un grupo de investigadores y posibles usuarios con el fin de poder definir las características principales con las que debería contar el equipo de supercómputo de acuerdo con las necesidades planteadas.

Además del trabajo coordinado y colegiado al interior de la Universidad, fue de gran apoyo contar con la asesoría de expertos de otros centros de supercómputo de México, siendo de gran apoyo las recomendaciones realizadas para la especificación a detalle de las bases, además de compartir recomendaciones para dar un valor agregado a la propuesta.

Implementación del equipo de cómputo

Uno de los principales retos que implica implementar un equipo de supercómputo es la habilitación del centro de datos; el espacio físico en donde se albergará la supercomputadora. En esta etapa trabajan muchas especialistas con el fin de cuidar aspectos de energía, humedad, conectividad y otros para cumplir con los estándares internacionales, y adaptarlos a las necesidades y características del espacio donde se implementará; por ejemplo, la altura del terreno para evitar inundaciones, el análisis de la temperatura para la selección del sistema de enfriamiento e incluso en nuestro caso buscar una solución antisísmica debido a que el lugar donde se colocó el equipo se considera una zona activa sísmicamente.

Implementación del proceso de operación

Con el objetivo de garantizar la operación del CADS y los servicios que éste ofrece, se estableció una serie de procesos y procedimientos que nos permitirán homologar la forma de trabajar, y establecer los puntos de control en cada uno de los servicios que se ofrecen, además de poder medir en un futuro los resultados obtenidos.

Los procesos se establecieron desde un inicio, de acuerdo con la forma que se identificó se debía trabajar en el CADS, pero una vez que se iniciaron actividades, estos se han tenido que ir adecuando a la necesidades del entorno, tanto para atender los proyectos de investigación, como las necesidades de capacitación y las actividades de vinculación.

Por otro lado, se tiene el reto de incorporar nuevo personal al CADS, ya sea por crecimiento del mismo o porque existe una gran movilidad, por lo que es fundamental contar con procesos, procedimientos y roles del personal para que nuestros servicios no se vean afectados.

Capacitación del personal

Uno de los objetivos principales dentro del CADS, es apoyar la generación de nuevo conocimiento y el desarrollo de proyectos de investigación relacionados con diversas áreas como Big Data; arquitectura computacional; computación distribuida; modelos de programación; ingeniería simulada; infraestructura de operaciones; matemáticas de gran escala; modelos moleculares; información cuántica; bioinformática; biomecánica; computación cognitiva; composición atmosférica; predicción climatológica; ciencias computacionales de la tierra; geofísica; energías; educación; análisis de comportamiento; simulación social, esto por mencionar algunas de las áreas de aplicación que se atienden dentro del Centro.

Cada una de las diferentes ramas de estudio demandan una capacitación específica para cada proyecto, es por lo anterior que contar con planes de capacitación constante se vuelve una de las líneas primordiales, tanto para el aprovechamiento del supercómputo como para la divulgación y generación del conocimiento adquirido en el CADS.

La capacitación se enfoca en dos áreas principales: una de ellas es para los usuarios

del equipo de supercómputo y otro para el personal que opera dicho equipo, el cual es primordial que cuente con capacitación continua en el manejo y operación de la infraestructura, además de capacitación en el manejo de diversas herramientas científicas o lenguajes de programación para atender los proyectos.

6. Impacto

En el CADS se apoya a diversas áreas del conocimiento y se permite colaborar de manera interdisciplinaria en pro de la generación de conocimiento y la formación de personal altamente capacitado en el manejo de infraestructura relacionada con la operación y mantenimiento de la supercomputadora.

Gracias a las supercomputadora de la Universidad de Guadalajara, se pueden apoyar proyectos relacionados con la ingeniería, medicina, geología, biología, astronomía, mecánica automotriz, entre otras, al obtener valiosos resultados mediante el procesamiento y almacenamiento de los datos a una mayor velocidad, logrando así simulaciones numéricas o visualizaciones científicas que antes no eran accesibles dentro de la institución.

Si bien, ya se contaba con un estudio de las necesidades de la Universidad con respecto a los servicios de supercómputo, una vez habilitado el espacio físico del CADS y poniendo en operación el equipo Leo Átrox, el acercamiento de la comunidad universitaria ha sido mayor al esperado. Por un lado, los estudiantes de maestría y doctorado empiezan a identificar al CADS como un espacio para capacitarse y apoyarse en sus proyectos de tesis; por otro lado, nuevos proyectos y áreas de aplicación han surgido por parte de los investigadores.

El reto hoy en día, es poder posicionarnos como un centro de excelencia en el occidente además de un lugar de referencia para el país, donde se apoyen proyectos de gran impacto, no solo en la ciencia, sino que sirvan tanto a la industria y gobierno que demandan soluciones a problemas complejos.

Agradecimientos

Este trabajo ha sido posible gracias al gran esfuerzo del equipo de la Coordinación General de Tecnologías de Información que con el apoyo de la alta dirección de la Universidad de Guadalajara, encabezada por la Rectoría General de esta casa de estudio y de su Coordinación General Administrativa, concretaron una visión clara de apoyo a la investigación y generación del conocimiento ante los nuevos paradigmas que ofrecen las tecnologías de información disponibles en el mercado.

Los autores desean expresar su agradecimiento al Rector General de la Universidad de Guadalajara en el período en que fue realizado el proyecto, el Mtro. Itzcóatl Tonatiuh Bravo Padilla, al Dr. Miguel Ángel Navarro Navarro, así como al Dr. Ricardo Villanueva Lomelí, actual Rector General de nuestra casa de estudio; así como a los entonces Coordinadores Generales Administrativos, la Dra. Carmen Enedina Rodríguez Armenta y el Mtro. Luis Alejandro León Dávila, pues sin su apoyo no hubiese sido posible esta implementación.

Referencias

1. Schwab, K, La cuarta revolución industrial, p.19. Editorial Debate, México, (2017)
2. Gutiérrez Díaz de León, LA, , La supercomputadora en la investigación, en Chinkes, Ernesto (Coord.), Potenciando la universidad del siglo XXI. Soluciones TIC para pensar la universidad del futuro, pp. 45-53, RedCLARA, (2017).
3. Idem.
4. Universidad de Guadalajara, Plan de Desarrollo Institucional (2014-2030), Universidad de Guadalajara (2014),

Integración de las de las plataformas (georreferenciación, recursos de investigación) para la Gestión de la de investigación: caso de la Universidad Nacional, Costa Rica

David Hine^{a,b}, dhine@una.ac.cr
<https://orcid.org/0000-0002-3469-9820>
Andrea MoraCampos^b, andrea.mora.campos@una.ac.cr,
<https://orcid.org/0000-0001-9813-2674>
Anayudel Gutiérrez Hernández^b, anayudel.gutierrez.hernandez@una.ac.cr
Universidad Nacional, Costa Rica
<https://orcid.org/0000-0002-2371-3666>

a Oficina de Transferencia de conocimiento y Vinculación Externa
b Vicerrectoría de Investigación

Resumen: El objetivo de esta ponencia es sistematizar el proceso de interconexión de plataformas para la divulgación de los datos de gestión de la investigación de la universidad. Entre estos sistemas están implementados con los siguientes *softwares*, CRIS, DSpace, VIVO y Georreferenciación (mediante Google Maps), según el marco de Ciencia Abierta. Se explica la importancia de la conexión y la divulgación de los datos, para rendición de cuentas, transparencia, insumo de toma de decisiones y trazabilidad de la investigación, y los procesos de cómo se llevó a cabo la interconexión de las plataformas. Finalmente, la universidad cuenta con una plataforma que permite el análisis de información con respecto a áreas de investigación, indicadores de investigación y publicaciones, espacios geográficos de acción del quehacer investigativo, lo que permite la trazabilidad de todos los elementos anteriores.

Palabras claves: CRIS, DSpace, VIVO, interconexión, georreferenciación.

Eje temático: Tecnologías en la innovación educativa e investigación.

Introducción, problemática y contexto

Las universidades públicas y las entidades gubernamentales de Costa Rica se encuentran en un momento histórico coyuntural donde es necesario desarrollar mecanismos de rendición de cuentas y transparencia sobre el uso de los fondos públicos, además de la divulgación de su quehacer académico en procura de fomentar la búsqueda de fondos para la inversión en investigación, acción social y docencia. Adicionalmente, los nuevos tiempos exigen una participación más activa por parte de las instituciones de educación superior en los procesos de la transferencia de conocimiento que generan hacia la sociedad, con especial énfasis en el sector productivo con miras a mejorar la empleabilidad y la especialización de las personas, en el contexto de la cuarta revolución industrial (González-Páramo, 2017).

Una estrategia solución a la situación anterior son las plataformas tecnológicas que disponen este tipo de información al público. Este tipo de herramientas permiten develar toda esta gestión de conocimiento y pueden contribuir con la instauración y el fortalecimiento de redes de cooperación entre comunidades científicas, académicas y de otra naturaleza.

Otro punto valioso de tener estas plataformas es que al liberar todo este tipo de información en la web, según estándares internacionales, le ha permitido a instituciones latinoamericanas como la nuestra, hacer visibles su quehacer intelectual a nivel mundial.

Ante esta situación la Universidad Nacional se encontró con la necesidad de implementar un mecanismo que le permitiera la gestión de los recursos de investigación, sin embargo la realidad de la universidad se asemejaba a los problemas que enfrentan las universidades e instituciones en estos temas, como el manejo de la difusión de los resultados de investigación, el currículo de las personas investigadoras; ya que todo lo anterior se basaba en iniciativas aisladas, de carácter nativo, y se prescinde de información, como el proceso de financiamiento de investigación, además del desconocimiento de la existencia de herramientas para gestionar la investigación (Grupo de Trabajo de Repositorios de REBIUN, 2013), redes de coautoría, disparadores de nuevas líneas de investigación, áreas geográficas de impacto, poblaciones beneficiarias, grupos meta, entre otros.

Por lo anterior expuesto, la universidad inició un proceso de integración y publicación de plataformas, con características especializadas e interoperables, donde la comunidad universitaria puede difundir su producción intelectual alrededor del mundo, lo que no solo los beneficia a ellos, sino a las universidades o instituciones patrocinadoras; esto a su vez se convierte en un proceso cíclico, puesto que la información difundida podría servir para resolver problemas a nivel nacional o internacional, fortalecer la innovación y la transparencia.

Esta liberación de información a través de estas plataformas puede contribuir a visibilizar la ciencia producida en América Latina, ciencia que ha sido históricamente invisibilizada por editoriales a las que no les ha interesado la producción de nuestra región, ya que al liberarse con estándares internacionales podrían difundir con cobertura internacional la información que publiquen y lograr la promoción de estas (Gamboa, 1996), dar visibilidad a las investigaciones científicas de calidad y acabar con el fenómeno de la ciencia perdida (REDALYC, 2011; Calvo, 2003; Monge-Nájera, Benavides-Varela y Morera, 2004; Garfield, 1984; Monge-Nájera, 2002).

Aunado a lo anterior, estos sistemas son herramientas de control de la productividad (Camí, Suñén-Piñol y Méndez-Vázquez, 2005); también, proveen la posibilidad de crear redes de información que, a través de Internet, fomenten el intercambio de datos regionales (Gamboa, 1996). Lo anterior, en resumen, podría fortalecer la calidad de las publicaciones e investigaciones.

Descripción de la solución tecnológica implementada

Alrededor del año 2010, en la universidad se propuso la construcción de un sistema que permitiera el registro de los proyectos de investigación, las personas responsables, participantes y el tipo de financiamiento. Sin embargo, con el crecimiento y demanda de datos de investigación para la toma de decisiones, se destacó la importancia de la evolución a un sistema CRIS (por sus siglas en inglés Current Research Information System), el cual, posteriormente, permitió gestionar los recursos de investigación por medio de la interoperabilidad con otras plataformas de la universidad; tales como la Red Académica Institucional (con el *software* VIVO), el Repositorio Académico Institucional (con el *software* DSpace) y el portal de Georreferenciación de Proyectos de Investigación mediante Google Maps, con el propósito de generar información para rendición de cuentas, transparencia y datos de Ciencia Abierta.

Lo siguiente que se realizó fue la ingesta de datos de forma automatizada en algunas plataformas relacionadas con Ciencia Abierta a partir del sistema CRIS de la universidad. Para una mejor comprensión del proceso, se definirá, a continuación, el alcance de cada plataforma.

VIVO (Red Académica)

Respecto, a la Red Académica de la Universidad Nacional, es una plataforma basada en el *software* VIVO desarrollado por la Universidad de Cornell, la cual ofrece un perfil para las personas investigadoras de la universidad. Esta permite visualizar la producción intelectual de personas autoras con afiliación institucional de la Universidad Nacional, así como las redes de autoría, estadísticas de la producción intelectual de la comunidad universitaria, según área de conocimiento, unidad académica, instituciones externas, áreas de impacto en la investigación, colaboraciones de investigaciones con otras universidades e instituciones, cantidad de publicaciones por unidades académicas, entre otros aspectos. Las personalizaciones que se le realizaron a esta plataforma fueron las siguientes: crear un identificador único para las personas con afiliación Universidad Nacional, agregar las áreas de conocimiento según la clasificación Frascati.

Para este proceso se realizó una recuperación de publicaciones por medio de las API, repositorios internos y un demo desarrollado para esos fines. Posteriormente, se hizo una importación en bloque utilizando la aplicación Karma para generar los modelos, aplicando las ontologías a los conjuntos de datos de acuerdo con las necesidades de la UNA (Mora & Hine, 2018).

Sistema de Georreferenciación

El quehacer investigativo de la Universidad Nacional es eje central para el desarrollo, el conocimiento y la cultura nacional; por lo tanto, la universidad decidió emprender la tarea de reunir los insumos necesarios para geolocalizar la investigación universitaria.

Georreferenciar es la técnica de posicionamiento global que le asigna a una entidad espacial una localización específica y única (ArcGIS, 2017). Esta condición de vinculación singular entre una entidad espacial (proyecto de investigación) y el territorio es un Sistema de Información Geográfica que se centra en laboratorios, programas, proyectos y actividades académicas vigentes.

El levantamiento de los geodatos se inició en octubre de 2017, se reunieron 395 coordenadas asociadas al territorio costarricense, hoy, en 2019, son 901 coordenadas las que han sido registradas, algunas de estas trascienden las fronteras. Cada punto representado en el mapa, tiene información asociada que permite un conocimiento general de este, los datos asociados son los siguientes: título de la investigación, código asociado al Sistema de Información Académica, facultad a la que pertenece, contacto de la persona responsable, población beneficiada, objetivo general, resumen y una fotografía que ilustra el quehacer académico.

Posteriormente, se exploraron diferentes opciones para la representación gráfica de los proyectos georreferenciados, después de analizar diferentes opciones, Google Maps en donde se encontraron muchas similitudes como la capacidad de realizar consultas y ejecutar *scripts*, sin embargo se eligió Googlemaps por razones como la facilidad de la utilización de las API que google tiene a disposición.

Proceso de interconexión entre sistema CRIS y el *software* VIVO

Las personas investigadoras realizan el registro de su proyecto de investigación en el CRIS con el fin de gestionar la información relacionada con este, a partir de este punto se implementó un mecanismo de extracción de dicha información para vincular las personas investigadoras con su producción intelectual en el *software* VIVO. Como producto de esta relación se obtuvo un currículo de las personas investigadoras que refleja la filiación, la producción intelectual, proyectos de investigación donde ha participado, áreas de investigación, entre otros elementos de importancia en la divulgación del quehacer académico.

Todo lo anterior ha permitido la trazabilidad de la investigación con elementos como, fuente de financiamiento, líneas de investigación, producción intelectual a partir de los proyectos, personas autoras y coautores, tanto internas como externas, volumen de producción por facultades y unidades académicas, todo esto brinda un conjunto de datos que favorecen la toma de decisiones informadas por parte las autoridades universitarias.

Sistema CRIS y el proyecto de Georreferenciación

A partir del proceso de georreferenciación supracitado, que consiste en agregar un dato que permita describir la ubicación geográfica de un proyecto y su área de trabajo o impacto en el sistema CRIS, se desarrolló una plataforma basada en Googlemaps que permitiera representar de forma gráfica dicha ubicación; además, incorpora un grupo de filtros que permite generar una serie de indicadores de forma dinámica. Algunos de los indicadores que se pueden extraer son, proyectos distribuidos por regiones socioeconómicas, por área de la ciencia, por regiones según índice de desarrollo humano, proyectos pertenecientes a facultades o unidades académicas determinadas entre otros.

Software VIVO y Georreferenciación

En la última fase de la propuesta se desarrolló el mapa interactivo con todos los proyectos georreferenciados que desarrolla la Universidad Nacional, que permite, dada una ubicación, obtener una ficha resumen con datos como, nombre del proyecto, persona investigadora responsable y poblaciones beneficiadas, además de un enlace a la Red Académica (VIVO), donde cada proyecto cuenta con un perfil completo, que le permite a la universidad relacionar y difundir toda la información relacionada con dicho proyecto como se muestra en la figura 1.



Fig. 1. Ficha de información de los proyectos georreferenciados.
Elaboración Propia apartir de <http://investigacion.una.ac.cr>

Sistema CRIS y DSpace

La interconexión del sistema CRIS con el DSpace permite relacionar la producción intelectual con los proyectos de investigación que les dieron origen. Esta relación se logró agregando a cada producción intelectual que se ingresa al repositorio un metadato de código de proyecto, este código de proyecto se obtiene de un consecutivo que genera el sistema CRIS para cada proyecto que se ingresa, el campo en DSpace se agrega utilizando, según la facilidad que incorpora Dspace, campos personalizados. Finalmente, esto permitió relacionar la producción intelectual de la universidad con los proyectos de investigación y por transitividad con las personas investigadoras participantes de estos.

Interconexión de las plataformas para la gestión de los recursos de investigación

La interconexión de todos los sistemas anteriores ha proporcionado un ecosistema como se muestra en la figura 2, el cual permite gestionar los recursos de investigación de forma clara con datos fiables, y un panorama más amplio que involucra diferentes variables, ya mencionadas, además de introducir la posibilidad de divulgar la información por medio de políticas de transparencia y rendición de cuentas.

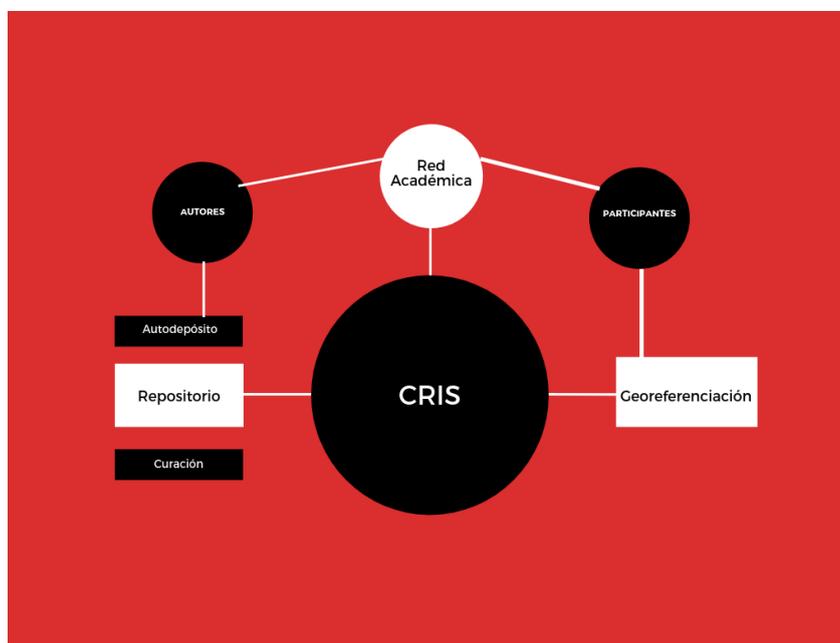


Fig. 2. Modelo de interconexión de las plataformas para la gestión de los recursos de investigación. Elaboración Propia.

Aspectos críticos y relevantes por resaltar

Ante la disyuntiva presente en muchas instituciones de tener desvinculados los sistemas, lo que ha provocado que se incurra en errores de calidad de la información, duplicidad de datos y de trabajo, surge la propuesta de hacer una investigación de los sistemas y cómo se podría hacer una vinculación de estos.

Algunos de los aspectos críticos que se han enfrentado a lo largo del proyecto fue la necesidad de concientizar sobre la importancia de una cultura de brindar en acceso abierto y gratuito, y de preferencia inmediato, la producción intelectual con la finalidad de difundir el quehacer de la Universidad Nacional, y los beneficios que se derivan de dicha práctica como captación de financiadores de proyectos de investigación, colaboraciones en investigaciones conjuntas con personas investigadores de instituciones en otras latitudes y convenios de cooperación.

Asimismo, se encontró cierta resistencia para poder acceder a los datos, aún cuando fueran de carácter público. Finalmente, encontramos inconsistencias en las colecciones de datos ante la ausencia de mecanismos de validación y curación.

Resultados obtenidos e impacto

Como resultado de dicha implementación, la universidad cuenta con un sistema integrado para la gestión y visualización de la producción intelectual de la Universidad Nacional. Incluso, se pueden visualizar integrantes y personas investigadoras, según sea el área de la ciencia, esto para generar redes internas y externas de investigación y otras áreas sustantivas y de acuerdo con el área de conocimiento de acuerdo con la clasificación de Frascati. Además, se cuenta con un identificador único para cada integrante de la comunidad universitaria que no sustituye iniciativas internacionales como el ORCID, pero si se empata en sistemas internos ante la ausencia de ORCID.

Esto ha impactado positivamente el quehacer académico, pues ha provocado que los autores de la universidad sean contactados como pares académicos, sean invitados a eventos científicos, sean evaluadores de proyectos de investigación en otros ámbitos. Por lo tanto, la integración puede generar un público y ampliar audiencias, lo cual resulta en fortalecer las redes académicas y productivas (S. Córdoba, conversación personal, 2017), adicionalmente, ha contribuido con autoridades en la toma de decisiones.

Ahora bien, el ejercicio de georreferenciación posibilitó que los equipos de investigación pudieran diversificar la forma de difundir su producción científica.

Se logra construir la cartografía completa que muestra el quehacer investigativo de la universidad, lo cual permite al equipo de investigación asociar información y datos para ampliar la forma en que se aborda el objeto de estudio, esto favorece trabajar e integrar información de distintas fuentes, escalas y datos. Este análisis puede hacerse de forma rápida, ágil y fiable.

El procesamiento de la información y la obtención de productos cartográficos es expedito con las herramientas que tiene la Universidad Nacional basadas en *software* libre.

Resultados de la interconexión de plataformas

Finalmente, la universidad cuenta con una plataforma que le permite entre otras cosas analizar la existencia de posible duplicidad de proyectos de investigación, análisis de los campos de investigación donde presenta fortalezas y debilidades, además, de poder generar indicadores que permitan la divulgación y la visibilización del trabajo realizado por la universidad, por otro lado, se pueden reconocer espacios geográficos del país que no han sido impactados o que podrían ser un nicho potencial en el campo de la investigación, todo esto relacionado con publicaciones académicas electrónicas ingestadas en el Repositorio Académico Institucional como productos de las actividades de investigación, docencia y acción social, lo que permite la trazabilidad entre personas investigadoras, proyectos, productos y fuentes de financiamiento.

Aprendizajes (aciertos y errores)

Necesidad de adoptar un identificador propio que fuera compatible con los sistemas que nos permitiera no depender de instancias externas por aquello de alguna eventualidad.

La necesidad de que los sistemas estén centralizados, pues la norma es que cada facultad maneje la información de sus proyectos y actividades académicas de forma local.

Este proceso ha estado lleno de retos no solo por el desafío técnico, sino también en el aspecto humano, en donde motivar a las personas encargadas de la gestión de proyectos a involucrarse de forma activa ha requerido igual tiempo y esfuerzo.

Los aciertos de este componente se centran, principalmente, en dos elementos: mantener sencillo el procedimiento de captura de geodatos y mostrar, de forma novedosa, las posibilidades de análisis de la información.

El primero ha reducido los errores en el registro asociado a las coordenadas y la resistencia de las personas a introducir información veraz. La estrategia de recurrir a aplicaciones naturalizadas entre las personas investigadoras, refuerza este acierto. El segundo es el reconocimiento del potencial de interpretación y traslape de la información disponible. Esto ha motivado a las autoridades universitarias a comprometerse con el proyecto y disponer del recurso que se requiera para conocer en profundidad el quehacer académico de la universidad.

Algunas de las experiencias negativas que podrían ser catalogados como errores fue que al inicio del proyecto el abordaje fue monolítico, centralizado en la instancia de tecnología, o en la parte política, sin tomar en cuenta a todos los actores involucrados, lo cual reflejó falencias en campos como la normalización de datos, consideraciones en licencias y derechos de autor, de tal manera que se desperdició el conocimiento de profesionales de diferentes campos, a lo interno de la institución, que pudieron haber resuelto en primera instancia.

"Divide y vencerás", en la universidad se adoptó la estrategia de desarrollar la propuesta desagregada por facultades, donde se vincula a las autoridades y personas tomadoras de decisiones, en conjunto con el personal de apoyo a la academia, de esta forma el desarrollo se hizo de forma incremental, y se llevó a cada facultad los

aprendizajes y mejoras en la experiencia con la facultad anterior.

Para la Red Académica se realizó un proceso de recuperación de publicaciones por medio de las API de Scopus, Web of Science, repositorios internos y un crawler desarrollado para la recuperación de metadatos. (Mora & Hine, 2019).

Impactos no esperados

La solicitud de replicación del modelo desde Consejo Nacional de Rectores CONARE para todas las universidades públicas desde el seno de la Comisión de Vicerrectores de Investigación, además de la solicitud de otras instituciones gubernamentales para implementar esta tecnología.

Propuesta del Instituto Geográfico Nacional para crear un Nodo de geoinformación para la Universidad Nacional.

Revitalizar el desarrollo profesional de las personas bibliotecólogas, de tal manera que se les integra a la era digital y se transforman en un recurso esencial de apoyo a la academia.

Referencias

1. ArcGIS. (2017). ArcGIS Resousers. Obtenido de ArcGIS Resousers:
2. <http://resources.arcgis.com/es/help/getting-started/articles/026n0000000s000000.htm>
3. Camí, J., Suñén-Piñol, E. y Méndez-Vázquez, R. (2005). Mapa bibliométrico de España 1994-2002: biomedicina y ciencias de la salud. *Medicina Clínica*, 124, 93-101.
4. Gamboa, O. (1996) .*Hacia el Establecimiento del Índice Latinoamericano de Publicaciones Científicas. Latindex*. Recuperado de <http://www.dgbiblio.unam.mx/servicios/dgb/publicdgb/bole/fulltext/volI2/latindex.html>
5. Grupo de Trabajo de Repositorios de REBIUN. (2013). Sistemas CRIS y Repositorios Institucionales en las Universidades Españolas. Recuperado de <https://www.rebiun.org/sites/default/files/2017-11/https://www.rebiun.org/sites/default/files/2017-11/CRISyRepositorios2013.pdf>
6. González-Páramo, J. (2017). CUARTA REVOLUCIÓN INDUSTRIAL, EMPLEO Y ESTADO DE BIENESTAR . 28 de Mayo 2019, de Real Academia de las Ciencias Morales y políticas Sitio web: <http://www.racmyp.es/R/racmyp/docs/anales/A95/A95-7.pdf>
7. Monge-Nájera, J., C. Benavides-Varela & B. Morera. (2004). ¿Cuáles son las revistas, libros y personas más influyentes en la biología neotropical? *Revista de Biología Tropical*, 51 (1).
8. Andrea Mora-Campos, David Hine (Universidad Nacional de Costa Rica) Exploring Costa Rica scholarly records by interactive knowledge networks: case of implementation of VIVO software in Universidad Nacional, Costa Rica

9. Latindex. (2017). *Latindex*. Recuperado de www.latindex.org
10. Redalyc. (2011) *Uso legal*. Recuperado de <http://redalyc.uaemex.mx/redalyc/media/principal/servicios/juridica.html>
11. Repositorio Nacional de Costa Rica (2016). Recuperado de <http://kimuk.conare.ac.cr/>

Capítulo 8

Logrando universidades digitales seguras

Automatización en la seguridad del perímetro de la red mediante tecnologías open-source

Miguel Morandi¹, Alejandro Cuadra¹, Graciela Becci¹ Marcelo Gómez¹

¹ Universidad Nacional de San Juan, Av Libertador San Martín Oeste 1109, San Juan, J5400ARL
Argentina

morandi@unsj.edu.ar, acuadra@unsj.edu.ar, gbecci@unsj.edu.ar, mgomez@unsj.edu.ar

Resumen. La seguridad de una red de dominio público como la de la universidad, presenta desafíos para su defensa ante las amenazas que evolucionan en forma permanente hacia formas más sofisticadas. Existen en el mercado opciones de defensa de la red altamente eficientes, no obstante, para una institución pública es difícil mantener actualizada una infraestructura propietaria, debido a los altos costos del equipamiento y la dependencia del fabricante para las actualizaciones del firmware y software propietario. En el presente artículo se explora la automatización en el sistema de seguridad del perímetro de la red, es decir en el Router de borde. La implementación se basa en software de dominio público ejecutado en una plataforma de hardware estándar, lo que flexibiliza la configuración y actualización de la protección mediante scripts con expresiones regulares. Esta implementación acompañada por buenas prácticas de ruteo ha llevado a una significativa reducción de las amenazas y disrupción de la red. Se muestra un caso de uso donde la flexibilidad de los sistemas de dominio público da la posibilidad de mantener actualizada la infraestructura, y mediante scripting poder desarrollar una alternativa eficaz para la protección a la red

Palabras Clave: seguridad en redes de datos, software de ruteo, Router de borde

Eje temático: Logrando universidades digitales seguras.

1. Introducción

La red de datos de la UNSJ y el consiguiente tráfico están a cargo del IDECOM, quien administra y mantiene la red en los distintos establecimientos distribuidos geográficamente en la provincia de San Juan. La red está compuesta por un parque tecnológico heterogéneo multimarca, el cual brinda servicio de acceso a internet, campus virtual, correo electrónico, páginas web, recursos locales de almacenamiento, entre otros. Históricamente el IDECOM contaba con dos proveedores de servicio de acceso a Internet, RIU (TASA) y CLARO. En el año 2009 se gestionaron, ante LACNIC, recursos propios de Internet (ASN y Direcciones IPv4 e IPv6). A partir del año 2017 se canceló el servicio con CLARO y se contrató el servicio de ArSAT, sumados a la instalación de un nodo para el tráfico nacional a través de CABASE, alojando el IDECOM al primer nodo IXP (Internet eXchange Point) de San Juan, logrando de esta manera que proveedores locales y la UNSJ intercambien tráfico nacional. Con respecto al hardware, a esa fecha se contaba con un Router Cisco 2800, haciendo falta una actualización de equipamiento para afrontar el nuevo tráfico de la red, ya que la conexión a Internet paso de 100 Mbps a 800 Mbps, sumando todos los proveedores, incluido el tráfico en tiempo real o streaming de las videoconferencias.

Se abordaron diferentes soluciones para actualizar la administración de la red de la Universidad, entre ellas se siguió la estrategia que se describe en este artículo, consistente en actualización tecnológica en el marco de restricciones impuestas por el presupuesto de una Universidad pública nacional. Por lo que la consigna fue uso de hardware de bajo costo, actualización de hardware existente, y uso de software libre para el desarrollo de un firewall que se actualice automáticamente en forma periódica. Este upgrade era absolutamente necesario dado el nuevo incremento de tráfico de la red. La actualización tecnológica consistió en el uso de un servidor estándar dotado de mayor capacidad de procesamiento, placa de red, memoria RAM, administrado por software de dominio público como el sistema operativo de red VyOS. Actualmente se logran filtrar direcciones IPv4, ASN y dominios; se está trabajando en la implementación de la red IPV6 y sus respectivas normativas de seguridad.

Por eso este artículo se centra en la descripción del automatismo de la seguridad en el perímetro de la red que llevó a mantener actualizado el sistema de prevención de intrusión IPS en el Router de borde. La implementación se basa en software de dominio público y plataformas de hardware estándar, lo que flexibiliza la configuración y actualización de la protección mediante scripts con expresiones regulares, que en forma complementaria defienden la red del tráfico indeseado.

En la Sección 2 se citan trabajos relevantes relacionados con el desarrollo de sistemas de defensa de la red usando recursos de dominio público. En la Sección 3 se detallan las etapas seguidas en la implementación, en la Sección 4 se detalla la experiencia en la implementación del sistema de seguridad, en la sección 5 se discuten los resultados y aprendizajes de la implementación, y en la Sección 6 se presentan las conclusiones y sugerencias de mejoras.

2. Antecedentes

La eficiencia de los sistemas de detección y prevención de intrusión a la red depende de que estos sistemas estén actualizados para actuar ante amenazas que evolucionan constantemente hacia formas más sofisticadas. Es difícil mantener actualizada una infraestructura propietaria, debido a los altos costos del equipamiento y la dependencia del fabricante para las actualizaciones del firmware y software propietario, con el consiguiente riesgo de falta de soporte por obsolescencia del producto. Una alternativa cada vez más eficiente la ofrecen los sistemas basados en distribuciones de dominio público. En el caso particular del hardware de ruteo se reemplaza eficientemente mediante la implementación de un sistema operativo abierto de red ejecutado en una plataforma de hardware convencional, comportándose con todas las funcionalidades de un equipo propietario.

La Universidad pública en particular ha presentado casos de experiencia en el desarrollo de su propia infraestructura informática basada casi en su totalidad en hardware cuya actualización sea costeable y software libre. En el caso particular de los Routers la implementación con el sistema operativo de red VyOS, basado en Vyatta, tiene popularidad por las prestaciones que ofrece, fácil implementación y posibilidad de adaptarlo a las necesidades particulares de cada institución [1].

Las prestaciones de los Routers por software se pueden ampliar mediante el desarrollo de módulos de software, como es el caso del desarrollo de un driver de red wireless, no contemplado en el sistema operativo original de VyOS, con el objeto de agregar seguridad wireless a la red [2]. Como contrapartida, los Routers por software encuentran dificultad para proveer altas velocidades de tráfico debido al procesamiento de los paquetes de datos en el kernel del sistema operativo. No obstante, al tratarse de sistemas abiertos, este tipo de problemas encuentran solución en las comunidades abiertas, tal como las propuestas para mejorar la performance del kernel optimizando el uso de recursos como es la CPU y memoria [3], o mediante la ejecución de tareas en paralelo [4], [5], o buscando caminos alternativos al kernel stack [6].

Los sistemas de detección de amenazas juegan un papel importante en la protección de la red mediante la generación de alertas. No obstante el análisis de estas alertas es laborioso e ineficiente cuando es realizado por un humano, y la toma de decisiones preventivas llega muchas veces tarde cuando el ataque ya penetró la red.

El uso de sistemas libres permite la implementación y automatización de estrategias de protección mediante el desarrollo de módulos de software y scripts incluyendo expresiones regulares, adaptando el firewall para prevenir ataques a las diferentes capas del modelo OSI, como por ejemplo la capa de enlace de datos, tráfico en la red y capas superiores de aplicación y acceso a la web [7], [2].

Una de las ventajas del desarrollo de firewall mediante script es que es particular de la red donde se desarrolla, dificultando el escaneo externo, que es una típica amenaza que tienen las soluciones de renombre. Por otro lado la desventaja es la necesidad de contar con personal capacitado para el desarrollo, test y actualización de un firewall dedicado. Los scripts para adaptar firewalls ofrecen la flexibilidad de poder incluir en su estrategia métodos heurísticos y expresiones regulares, como por ejemplo para inspeccionar tráfico web [8], [9].

3. Etapas en la implementación

Para afrontar las amenazas a la red incrementadas con el aumento de tráfico de datos de la UNSJ, se propuso la estrategia de usar un sistema de seguridad en el perímetro de la red como un sistema de prevención de intrusión IPS en el Router de borde, actualizando la información mediante una plaqueta Raspberry Pi, según el esquema de la Figura 1.

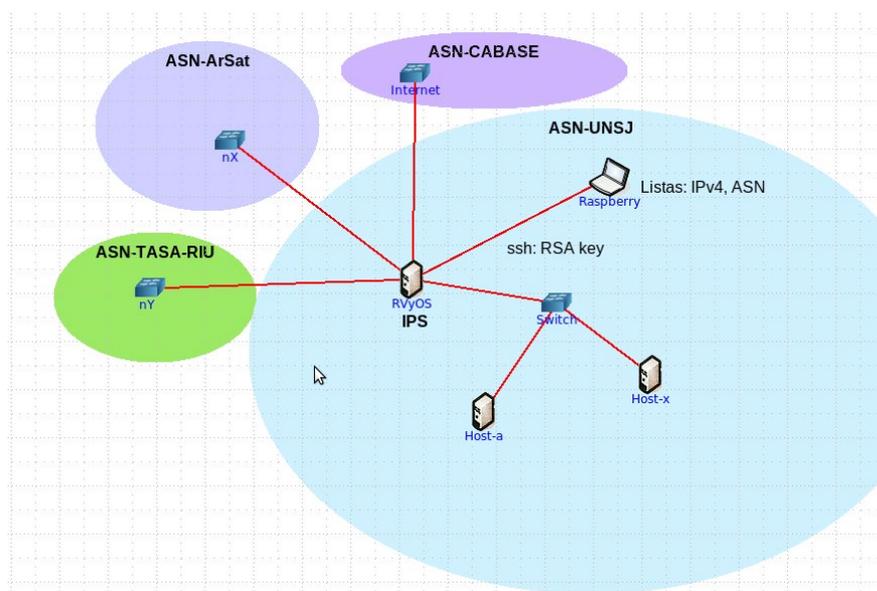


Fig. 1. Protección en el perímetro de la red: Sistema de Prevención de Intrusión IPS, donde el Router VyOS es el encargado de enviar el tráfico filtrado a la red interna de la Universidad, mientras que la actualización del sistema está a cargo de la placa Raspberry Pi.

El Router de borde se conformó mediante un sistema operativo de red de dominio público y la actualización tecnológica con hardware de bajo costo, para el desarrollo de los firewalls descritos a continuación.

3.1 Adopción de tecnología open-source

A comienzos de 2018, se decidió reutilizar un servidor obsoleto para estos fines, HP Proliant DL360 de 64 bits, y convertirlo en Router. La configuración del Router significó un incremento en la capacidad de procesamiento mediante un procesador 32/64 2GHz Xeon 4 núcleos, se actualizó la placa de red a 2 Gbps y 4 Gbps (modo switch), y se incrementó la memoria RAM a 2048 MB. Con estas mejoras la plataforma quedó lista para ser convertida en Router, con la ventaja de poder realizar mejoras a futuro sin depender de licencias de marcas.

Implementación del Router por software

Con respecto al software del router se evaluó las ventajas comparativas entre implementar un sistema operativo de red y un firewall, implementados en una plataforma de hardware estándar. Por su amplia difusión se compararon pfSense y VyOS. Si bien pfSense es un firewall de red de distribución libre y de muy fácil implementación [10], en este caso para el desarrollo de un sistema de prevención de intrusión se prefirió el análisis de otro tipo de herramienta más completa para ejecutarse en un Router de borde. Por eso se eligió VyOS, que es un sistema operativo de red que soporta protocolos de ruteo BGP, OSPF y lenguajes para implementar políticas de ruteo complejas. También incluye ruteo real y virtual, firewall y VPN. Con respecto a la implementación, VyOS tiene la particularidad de poder ejecutarse en una amplia variedad de plataformas físicas y virtuales, tanto en pequeñas placas x86 como en grandes servidores, en Virtual-Machines, y Kernel-based VM, en hypervisors Hyper-V y Xen. La administración de VyOS se realiza mediante interface de línea de co-mandos similar a los routers por hardware, y estos comandos pueden incluirse en un script para ser ejecutados en conjunto en forma automática. También permite mantener el estado de las configuraciones, archivarlas y eventualmente volver a versiones anteriores. Todas estas características fueron decisivas a la hora de invertir recursos en la curva de aprendizaje de VyOS, ganando en el manejo de una herramienta potente para la administración del tráfico de la red. Se instaló la versión VyOS 1.1.7 de 32 bits, se realizó la migración completa de la configuración del router Cisco al router VyOS.

Para agilizar el monitoreo del estado de la red de la UNSJ distribuida geográficamente en distintos departamentos, se instaló una aplicación con interface gráfica que permite el monitoreo en tiempo real de los enlaces y recursos de equipamientos críticos de la red (router, switch raíz, UPS, etc). Esta herramienta, que además tiene la capacidad de enviar alarmas por mensajes de texto, permitió encontrar cuellos de botella en varios segmentos y poder solucionarlos, logrando que en la actualidad no exista saturación de los enlaces.

3.2 Pasos del Sistema de Prevención de Intrusión a la red

Para prevenir el tráfico indeseado, se desarrolló un Sistema de Prevención de Intrusión a la red (IPS), abordando tres distintos frentes: bloqueo por hosts y redes IPv4 en listas negras, bloqueo de tráfico anunciado por sistemas autónomos en listas negras, que son identificados por el número del sistema autónomo ASN, todos con antecedentes de phishing, malware, fraude, ataques, virus, etc. Como fuente del filtrado se buscaron otras listas además de la usada por ntop, que fueran de acceso al dominio público, dado el costo de los servicios pagos.

Raspberry Pi en la seguridad del perímetro de la red

Como parte de la seguridad del perímetro de la red se usó una placa Raspberry Pi 1, con sistema operativo Raspbian usando la conexión Ethernet para la comunicación con el switch VyOS. Para que el Sistema de Prevención de Intrusión IPS funcione de forma

autónoma, se desarrolló un mecanismo para la automatización de la búsqueda de posibles amenazas a la red para actualización periódica del firewall, consistente en los siguientes pasos.

Paso 1: Ejecución de script de acondicionamiento de listas negras. En una placa Raspberry Pi se ejecuta el script que extrae automáticamente de las listas negras la información para clasificarla por host, network y ASN, etc., y limpia caracteres de comentarios, acondiciona los registros de la BD de spamhaus i.e.

Script en Raspberry Pi:

1. Descarga de listas negras de dominio público publicadas en Internet
2. Limpia el contenido de caracteres y comentarios para obtener las direcciones IP puras
3. Separa la información en 3 archivos usando expresiones regulares: direcciones de host, de red (IPv4), y ASN (IPv4-v6).
4. Agrega comandos VyOS

Paso 2: Envío del archivo ejecutable al Router VyOS. La placa Raspberry Pi se comunica con Router VyOS via ssh usando RSA Key para el envío de los archivos.

Paso 3: Ejecución en VyOS del filtro.sh y configuración para la ejecución automática del archivo.

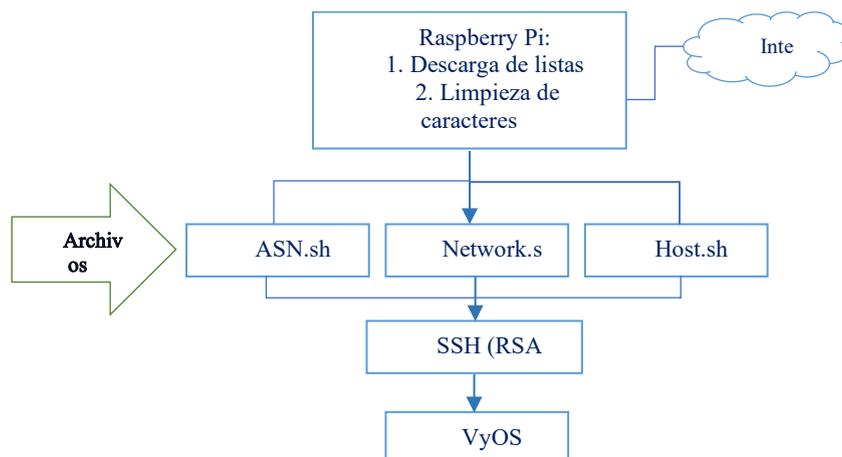


Fig. 2. Raspberry Pi en la seguridad del perímetro de la red. Pasos para la automatización del sistema de prevención de intrusión, acciones ejecutadas en la placa Raspberry Pi y su conexión con el Router VyOS.

Como acción complementaria de seguridad, la placa Raspberry Pi fue configurada para acceder a un grupo de VLANs troncalizadas en la conexión con el Router VyOS, esto permitió realizar el monitorear del tráfico de las VLANs aguas abajo del Router.

3.3 Seguridad dentro del perímetro de la red

La seguridad interna se reforzó mediante un Sistema de Detección de Intrusión

(IDS), se usó ntop, con licencia para instituciones educativas, para el monitoreo del tráfico de la red. Se instaló ntopng, analizador de tráfico, y nProbe para inspección de contenido de aplicaciones en la capa 7, los cuales son capaces de analizar mediante sFlow los puertos del Router y del switch de core. Ntop recibe resúmenes de tráfico sFlow provenientes del router VyOS, que son resúmenes de paquetes de datos truncados a nivel de capa 2 de enlace, sin payload con el propósito de monitoreo de la red. Esto le permite tener un buen panorama de lo que sucede en la red. Compara este tráfico con una lista pública de hosts y networks IPv4 maliciosas, elaborada por emergingthreads.net en base hosts y redes completas comprometidas por diferentes motivos de abuso, spam, ataque, etc. Cuando ntop detecta tráfico indeseado genera una alerta. En esta implementación se generaron inicialmente entre 10000 y 15000 alertas diarias, lo cual era un considerable número de posibles amenazas.

Si bien ntop genera alertas, no bloquea el tráfico por lo que es necesario analizar las alertas para tomar acción. Para ello se guarda la información en logs y se analizan los casos particulares, esto es posible hacerlo con intervención humana dada la drástica reducción a alrededor de 10 casos diarios. No obstante, queda a futuro automatizar el proceso de actualización diaria del Router.

4. Experiencia de la implementación

4.1 Implementación del IPS: bloqueo de tráfico de hosts y redes

Primero se descargaron los archivos de texto de las listas negras, para luego extraer las direcciones IPv4 de host y de redes (grupos de hosts). Para la búsqueda de direcciones IPv4 de host, se usó el comando de búsqueda grep aplicado al archivo de texto, más la expresión regular quedando el comando:

```
grep -E -o "([0-9]{1,3}[\.] ){3}[0-9]{1,3}" archivo.txt
```

Si bien esta expresión regular no detecta direcciones IPv4 no válidas, como por ejemplo 256.256.256.1, se asumió que todas las direcciones del archivo de texto están conformadas y son válidas según el protocolo de direccionamiento.

Procedimiento similar se siguió para la detección de redes, cambiando solamente la expresión regular que acompaña a la búsqueda mediante el comando grep, quedando en:

```
"([0-9]{1,3}[\.] ){3}[0-9]{1,3}\/[0-9]{1,2}"
```

Igualmente se asume que todas las direcciones IPv4 y la descripción de la red son válidas y cumplen con el protocolo de direccionamiento. Con estas direcciones de host y redes se crea un nuevo archivo que contiene solamente direcciones IP para configurar el firewall según el paso siguiente.

Tercer paso, agregar comandos VyOS. Debido a que la ejecución de scripts de

migración es más simple que alterar los archivos de configuración de VyOS, es que se opta por agregar los comandos VyOS a las direcciones IP para que el script pueda ser ejecutado en el Router. Por ejemplo, se usó el comando sed para agregar la información necesaria para ingresar cada host en una lista de host, quedando el comando sed:

```
sed -i -e 's/^/set firewall group address-group malware-hosts address /'
```

Con este comando se agrega la información al comienzo de cada línea. El resto del formateo del archivo se hace mediante la ayuda de los comandos cat y echo para ejecutar los comandos VyOS. A continuación, se muestra parte del script con las acciones mencionadas:

```
cd /tmp
rm malware.sh
rm emerging-Block-IPs.txt
wget https://rules.emergingthreats.net/fwrules/emerging-Block-IPs.txt
#hosts
echo '#!/bin/vbash' >> malware.sh
echo 'source /opt/vyatta/etc/functions/script-template' >> malware.sh
echo 'configure' >> malware.sh
#echo 'del firewall name 152 rule 18' >> malware.sh
#echo 'del firewall name 152 rule 19' >> malware.sh
#echo 'del firewall name 152 rule 18' >> malware.sh
#echo 'del firewall name 152 rule 19' >> malware.sh
#echo 'del firewall group address-group malware-hosts' >> malware.sh
#echo 'del firewall group network-group malware-networks' >> malware.sh
grep -E -o "([0-9]{1,3}[\.]){3}[0-9]{1,3}" emerging-Block-IPs.txt > malware.tmp
sed -i -e 's/^/set firewall group address-group malware-hosts address /' malware.tmp
cat malware.tmp >> malware.sh
rm malware.tmp
#networks
grep -E -o "([0-9]{1,3}[\.]){3}[0-9]{1,3}\v[0-9]{1,2}" emerging-Block-IPs.txt > malware.tmp
sed -i -e 's/^/set firewall group network-group malware-networks network /' malware.tmp
cat malware.tmp >> malware.sh
rm malware.tmp
echo "commit" >> malware.sh
echo "save" >> malware.sh
echo "exit" >> malware.sh
```

El resultado es un archivo con extensión .sh listo para ser ejecutado en el router VyOS. Lo que resta por hacer es enviar el archivo al router y ejecutarlo. El archivo se envía al Router mediante el comando scp:

```
scp -2 -i ~/.ssh/org_rsa /tmp/malware.sh
vyos@172.16.254.1: /tmp/malware.sh
```

El archivo se ejecuta en el Router mediante el protocolo Secure Shell usando el comando ssh:

```
ssh -2 -i ~/.ssh/org_rsa vyos@172.16.254.1 /tmp/./malware.sh
```

Este script se ejecuta en el Router VyOS detectando las direcciones IPs. El último

paso es automatizar su ejecución, para ello usamos el comando crontab para ejecutar el script a determinadas horas.

4.2 Bloqueo de tráfico anunciado por Sistemas Autónomos ASN

El objetivo es bloquear tráfico IPv4 anunciado por Sistemas Autónomos ASN reportados en listas negras, de forma tal que el firewall se actualice automáticamente. Similarmente a lo indicado para tráfico de hosts y redes, se usaron listas negras de ASN (sistemas autónomos de BGP) de dominio público, en particular las provistas por el sitio spamhaus.org (<https://www.spamhaus.org/drop/asndrop.txt>)

Dicha lista se encuentra en texto plano, con comentarios (iniciados con ;) en donde en cada línea se lista un ASN, por lo que se siguió el procedimiento similar al anterior, limpiando las listas de comentarios para obtener los ASN puros. El comando grep más la expresión regular es:

```
grep -E -o "(AS)[0-9]{1,6}" archivo.txt
```

Esta expresión regular "(AS)[0-9]{1,6}" sólo deja pasar a los sistemas autónomos (ASN).

El próximo paso es agregar los comandos VyOS para que el script pueda ser ejecutado en el router. Para dar formato al archivo de texto puro usamos el comando sed, de manera similar a como se hizo con las IPs. El resto del formateo del archivo se hace mediante la ayuda de los comandos cat y echo. El resultado es un archivo con extensión .sh listo para ser usado en el router VyOS. Lo que resta por hacer es enviar el archivo al router y ejecutarlo. Para enviar el archivo usamos el comando scp:

```
scp -2 -i ~/.ssh/org_rsa /tmp/asndrop.sh  
vyos@172.16.254.1:/tmp/asndrop.sh
```

El archivo se ejecuta en el Router mediante el protocolo Secure Shell usando el comando ssh:

```
ssh -2 -i ~/.ssh/org_rsa vyos@172.16.254.1 /tmp/./ asndrop.sh
```

El script actualiza las IPs en el router VyOS. El último paso es automatizar su ejecución mediante el comando crontab.

5. Resultados y aprendizaje en la implementación

Las mejoras significaron un incremento de la velocidad de la red de 2 Gbps a 4 Gbps, y se pudo implementar firewall con registro de estado (stateful) mediante IPTables a cambio de Listas de Acceso. Además, la administración de la red mediante scripting en el Shell de Linux (Bash Shell) aportó flexibilidad para implementar desarrollos de control de tráfico. Tras realizar la migración mejoró la calidad del servicio, especialmente del tráfico en tiempo real, con lo que se eliminaron los problemas en las

videoconferencias, disminuyendo la latencia de respuesta a 24ms comparada con los 30 ms que tenía anteriormente. El porcentaje de uso de recursos de CPU y RAM disminuyó con la implementación de VyOS, llegando a un 10% de uso de CPU y 12% de RAM.

El incremento de la tasa de transmisión se debe fundamentalmente a un upgrade en el hardware de base del Router. Por el momento no se han realizado pruebas para detectar retardos debido a procesos en el kernel [4]. En caso de detectarse latencias, y al tratarse de software libre, es factible compensar las demoras en el kernel mediante la implementación de estrategias específicas, algunas de las cuales fueron mencionadas en la Sección 2.

Los resultados de aplicar este sistema de seguridad en el perímetro de la red son mostrados en la Fig 3. Las capturas de pantalla de la Fig 3 a) muestra el efecto de eliminar como prueba una de las reglas del firewall. Como resultado se observaron 5278 intentos de intrusión a la red, de los cuales se detectó una dirección IP origen del mayor número de intentos.

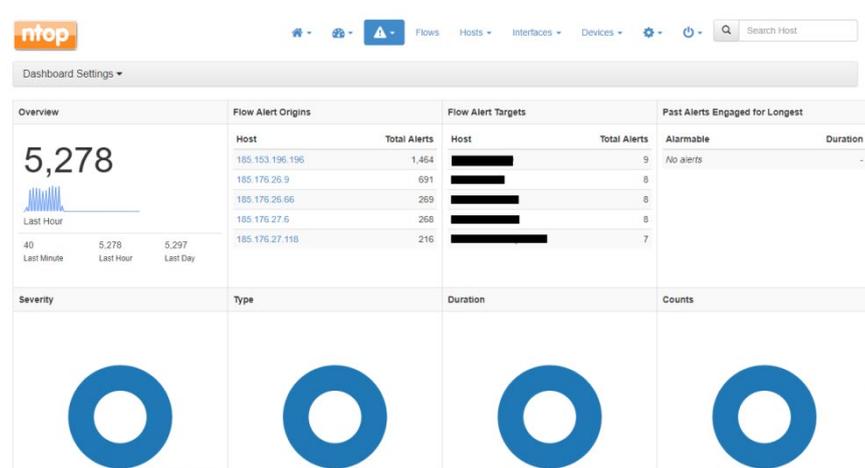


Fig 3. a) Descripción de los 5278 intentos de intrusión a la red al suspender como prueba la regla del firewall.

La Fig 3 b) da mayores detalles de las direcciones IP atacantes, tales como dirección y puerto de origen y destino, protocolo usado, etc. Para evaluar si realmente se trata de una intrusión verdadera se realizó una búsqueda de los datos de esta IP y se comprobó que, en el caso de no ser un spoofing de dirección, es una dirección que no tiene relación con las funciones académicas de la Universidad.

The screenshot shows the ntop interface with the 'Flow Alerts' tab selected. The table below represents the data shown in the interface:

Date/Time	Severity	Alert Type	Description	Actions
Tue Apr 23 21:16:12 2019	Error	! Blacklisted Flow	Client, server or domain is blacklisted [Flow: 185.153.196.196 @ 56667 → [REDACTED] 3382] [L4 Protocol: TCP]	Exploit Delete
Tue Apr 23 21:16:12 2019	Error	! Blacklisted Flow	Client, server or domain is blacklisted [Flow: 185.176.26.9 @ 54519 → [REDACTED] 33853] [L4 Protocol: TCP]	Exploit Delete
Tue Apr 23 21:16:12 2019	Error	! Blacklisted Flow	Client, server or domain is blacklisted [Flow: 185.176.26.9 @ 54519 → [REDACTED] 6369] [L4 Protocol: TCP]	Exploit Delete
Tue Apr 23 21:16:12 2019	Error	! Blacklisted Flow	Client, server or domain is blacklisted [Flow: 185.176.27.70 @ 47168 → [REDACTED] 1322] [L4 Protocol: TCP]	Exploit Delete
Tue Apr 23 21:16:12 2019	Error	! Blacklisted Flow	Client, server or domain is blacklisted [Flow: 185.153.196.196 @ 56667 → [REDACTED] 33853] [L4 Protocol: TCP]	Exploit Delete
Tue Apr 23 21:16:12 2019	Error	! Blacklisted Flow	Client, server or domain is blacklisted [Flow: 185.254.122.17 @ 43696 → [REDACTED] 24024] [L4 Protocol: TCP]	Exploit Delete
Tue Apr 23 21:16:12 2019	Error	! Blacklisted Flow	Client, server or domain is blacklisted [Flow: 185.153.196.196 @ 56667 → [REDACTED] 3398] [L4 Protocol: TCP]	Exploit Delete
Tue Apr 23 21:16:08 2019	Error	! Blacklisted Flow	Client, server or domain is blacklisted [Flow: 185.176.27.38 @ 57857 → [REDACTED] 48585] [L4 Protocol: TCP]	Exploit Delete
Tue Apr 23 21:16:08 2019	Error	! Blacklisted Flow	Client, server or domain is blacklisted [Flow: 185.153.196.196 @ 56667 → [REDACTED] 3372] [L4 Protocol: TCP]	Exploit Delete
Tue Apr 23 21:16:08 2019	Error	! Blacklisted Flow	Client, server or domain is blacklisted [Flow: 185.153.196.196 @ 56667 → [REDACTED] 3372] [L4 Protocol: TCP]	Exploit Delete

Fig 3. b) Descripción de los 5278 intentos de intrusión a la red con detalles de direcciones IP de origen, puerto y protocolo usado.

La placa Raspberry Pi ha demostrado muy buena performance y robustez en su desempeño dentro del sistema de seguridad de la red. Tanto el hardware como el software están preparados para trabajar en forma continua y confiable. Al principio de la implementación se detectaron resets aleatorios de la placa, luego de un análisis pormenorizado se llegó a la conclusión que en la asignación de VLANs para el análisis de tráfico, todas estaban asociadas automáticamente a una sola dirección MAC. Esto se cambió, asignando direcciones MAC a cada VLAN y el problema del reset se solucionó en forma definitiva.

6. Conclusión y sugerencias de mejoras

La implementación de este sistema de prevención de intrusión basado en script y facilitado por el uso de un Router por software ejecutado en una plataforma genérica demuestra una vez más que los sistemas de dominio público son eficientes y flexibles para solucionar amenazas a la red, que evolucionan dinámicamente.

La estrategia del sistema de prevención de intrusión presentado acompañado de buenas prácticas de ruteo ha reducido significativamente el nivel de ataques e incidentes en la red. No obstante siguen existiendo casos de actividad contra la red no detectados por ambos sistemas, lo que prueba que es necesario el análisis ulterior y la implementación de estrategias complementarias para que el sistema reaccione en forma apropiada cuando se produzca una intrusión.

La versión de VyOS usada es 1.1.7 que no soporta la implementación del sistema RPKI (Resource Public Key Infrastructure). Dado que la Universidad está comprometida con la administración responsable de la red y la mejora del ruteo global, es que se están siguiendo los lineamientos dictados por MANRS (Mutually Agreed Norms for Routing Security) para los operadores de red, por lo que en este caso también

incluye la actualización del router para soportar la implementación del sistema RPKI.

Como próximas tareas en relación a la performance de la red, queda determinar la incidencia del Router por software en la tasa de transmisión, y en caso de notar una ralentización, implementar algún mecanismo para sortear el cuello de botella. Con respecto a la eficacia del firewall es importante determinar si existen falsos positivos que puedan estar afectando negativamente a la red de la Universidad.

Referencias

- [1] Del Brocco, A.D (last), "Gestión de Servicios Informáticos con Software libre Universidad Nacional de Quilmes," *Jorn. Argent. Softw. Libre*, pp. 30–39, 2013.
- [2] S. Chaitra and R. Sharma, "Integration of Software Router with Wi-Fi for Enhanced Security," in *2017 IEEE 7th International Advance Computing Conference (IACC)*, 2017, pp. 33–36.
- [3] C. Hong, K. Lee, J. Hwang, H. Park, and C. Yoo, "Kafe: Can OS Kernels Forward Packets Fast Enough for Software Routers?," *IEEEACM Trans. Netw.*, vol. 26, no. 6, pp. 2734–2747, Dec. 2018.
- [4] D. Cerović, V. Del Piccolo, A. Amamou, K. Haddadou, and G. Pujolle, "Fast Packet Processing: A Survey," *IEEE Commun. Surv. Tutor.*, vol. 20, no. 4, pp. 3645–3676, Fourthquarter 2018.
- [5] S. Gallenmüller, P. Emmerich, R. Schönberger, D. Raumer, and G. Carle, "Building Fast but Flexible Software Routers," in *2017 ACM/IEEE Symposium on Architectures for Networking and Communications Systems (ANCS)*, 2017, pp. 101–102.
- [6] Z. Li, "HPSRouter: A high performance software router based on DPDK," in *2018 20th International Conference on Advanced Communication Technology (ICACT)*, 2018, pp. 503–506.
- [7] M. R. Zalbina, T. W. Septian, D. Stiawan, M. Y. Idris, A. Heryanto, and R. Budiarto, "Payload recognition and detection of Cross Site Scripting attack," in *2017 2nd International Conference on Anti-Cyber Crimes (ICACC)*, 2017, pp. 172–176.
- [8] I. Mukhopadhyay, K. S. Gupta, D. Sen, and P. Gupta, "Heuristic Intrusion Detection and Prevention System," in *2015 International Conference and Workshop on Computing and Communication (IEMCON)*, 2015, pp. 1–7.
- [9] B. Chaudhari, P. Gothankar, A. Iyer, and D. D. Ambawade, "Wireless network security using dynamic rule generation of firewall," in *2012 International Conference on Communication, Information Computing Technology (ICCICT)*, 2012, pp. 1–4.
- [10] "Getting Started With pfSense Software." [Online]. Available: <https://www.pfsense.org/getting-started/>. [Accessed: 20-Mar-2019].

Establecimiento del Plan de Continuidad del Centro de Operaciones de Videoconferencia de la DGTIC UNAM

José Luis Rodríguez Valdez

Universidad Nacional Autónoma de México
Centro de Operaciones de Videoconferencia (VNOG UNAM)
Círculo Exterior. S/N. Ciudad Universitaria, Ciudad de México
luisfca@unam.mx

Resumen. En el presente trabajo se presentan las actividades realizadas para establecer un Plan de Continuidad del Centro de Operaciones de Videoconferencia (VNOG, por sus siglas en inglés) de la Dirección General de Cómputo y de Tecnologías de la Información y Comunicación (DGTIC) de la Universidad Nacional Autónoma de México (UNAM). Se describe la operación de los equipos y sistemas del VNOG UNAM, la importancia de los servicios de comunicación audiovisual interactiva o videoconferencia en la UNAM, y la metodología utilizada para elaborar el Plan de Continuidad. De manera representativa en este trabajo, se describe la identificación de los usuarios de los servicios de videoconferencia, y el análisis de riesgos de los activos críticos, físicos y lógicos de la infraestructura central. De igual forma, se presentan a manera de ejemplo, algunas de las actividades para el restablecimiento de los servicios de videoconferencia que han permitido al personal operativo de TI de la dependencia identificar los recursos disponibles para el restablecimiento de los servicios del VNOG UNAM en caso de incidentes con el Manual de Contingencias y Recuperación de Desastres.

Palabras Clave: seguridad informática, plan continuidad, videoconferencia, VNOG.

Eje temático: Logrando universidades digitales seguras / Gobierno de TI.

1. Introducción

La infraestructura tecnológica de videoconferencia H.323 de la Dirección de Telecomunicaciones de la Dirección General de Cómputo y de Tecnologías de la Información y Comunicación (DGTIC) [1], brinda conectividad a los servicios de educación a distancia como cursos, exámenes profesionales, seminarios entre las dependencias de la Universidad Nacional Autónoma de México (UNAM) o entre instituciones educativas en México y el mundo. Y, por otra parte, permite la realización de reuniones de trabajo para la realización de proyectos académicos, así como reuniones de alto nivel entre las distintas entidades de la institución para informes de actividades o la toma de decisiones importantes de la administración universitaria. En su conjunto, se brindan un promedio de 1,100 sesiones de videoconferencia multipunto por año. Es así como los servicios de videoconferencia no solo ha sido un medio de comunicación educativo, sino también estratégico en casos de contingencias institucionales, tales como los paros de labores por parte del personal sindicalizado, y en situaciones extraordinarias que son ajenas a la actividad académica.

1.1 EL VNOC UNAM

El Centro de Operaciones de Videoconferencia (VNOC, por sus siglas en inglés) de la UNAM es parte de la Dirección de Telecomunicaciones de la DGTIC y tiene como principal objetivo facilitar el acceso a recursos de comunicaciones audiovisuales, así como proporcionar los servicios de videoconferencia y tecnología educativa a la comunidad universitaria a través de la Red de Videoconferencia de la UNAM (RVUNAM) y la Red Nacional de Videoconferencia conformada por las instituciones de afiliadas a la Red Nacional de Investigación y Educación (RNIE) en México.

Los servicios de videoconferencia multipunto son efectuados por diversos sistemas que brindan información del servicio, efectúan solicitudes en línea, y equipos que realizan la conexión simultánea de terminales de videoconferencia. Estas terminales pueden ser equipos de sala (H.323/SIP), o equipos personales (computadoras) y móviles (laptops, tabletas electrónicas, teléfonos inteligentes, etc.) que se encuentran dentro y fuera de las redes de datos universitarias. De acuerdo con *Wainhouse Research* [2], empresa independiente de investigación de mercados orientada a temas de comunicaciones multimedia y conferencias; un VNOC es una forma especial de un NOC (*Network Operations Center*) que está enfocado en el monitoreo y administración de colaboración visual y videoconferencia. Desde una perspectiva *top-level*, el NOC y el VNOC comparten los mismos requerimientos funcionales como:

- Monitoreo en tiempo real en un ambiente de producción.
- Detección inicial de problemas.
- Documentación y seguimiento de todos los aspectos operacionales de un usuario.
- Solución básica de problemas (*troubleshooting*).
- Coordinación y administración de otros recursos, internos y externos.
- Escalamiento de problemas a otros niveles si es requerido.

El VNOC UNAM está dedicado a las actividades de los usuarios de la Red de Videoconferencia de la UNAM (RVUNAM), y adicionalmente brinda apoyo técnico a miembros de otras redes académicas conectadas a su red como la Red de

Videoconferencia de CUDI (RVCUDI), que son las sedes de instituciones asociadas y/o afiliadas a la Corporación Universitaria para el Desarrollo de Internet A. C. (CUDI) en México, y otras Redes Nacionales de Educación e Investigación (RNEI), que son los usuarios y sedes en instituciones asociadas o afiliadas a una Red Nacional de Educación e Investigación con la cual CUDI tenga convenio de colaboración.

1.2 Operación general del VNOC UNAM

La RVUNAM está integrada de cerca de 350 sedes³⁰ que cuentan con equipo de sala H.323. Los responsables técnicos de los servicios audiovisuales en las dependencias universitarias que requieren de una conferencia multipunto, son los que realizan la solicitud desde la página del VNOC (<https://vnoc.unam.mx>). Este proceso es realizado en la plataforma de *Google Forms*, en la que se les solicita a los usuarios la fecha, horario y nombre de la sesión. Posteriormente, se envía la solicitud vía email para agendar el evento en la plataforma de calendario de Google (*Google Calendar*). Es aquí donde el personal de reservaciones verifica las necesidades del usuario, principalmente si requiere conectividad a sistemas de sala y/o equipos móviles, o *streaming*, y asigna los recursos de videoconferencia multipunto disponibles. En la fecha y hora del evento, se realiza el monitoreo de las sesiones desde las consolas de administración de los equipos multipunto y de forma audiovisual desde otros equipos de videoconferencia dispuestos en el lugar de trabajo para apoyar cualquier incidente a los usuarios.

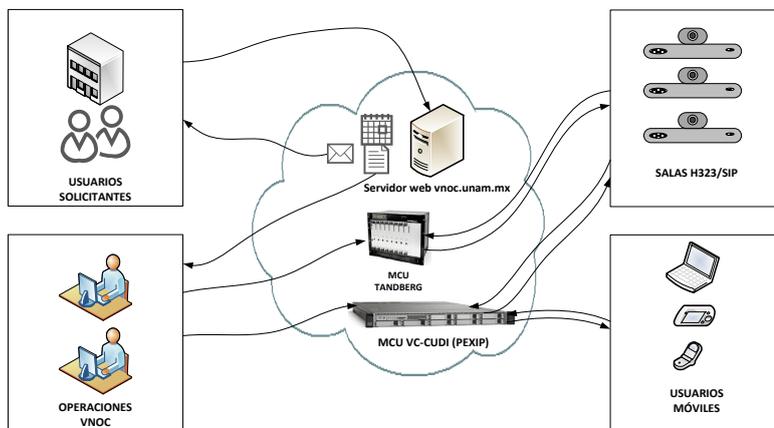


Figura 1. Esquema de operación de servicios del Centro de Operaciones de Videoconferencia. Fuente: realización propia.

1.3 Necesidad y justificación del Plan de Continuidad del VNOC

El VNOC UNAM presta los siguientes servicios a la comunidad universitaria [3]:

³⁰ Salas de Videoconferencia UNAM, <https://vnoc.unam.mx/index.php/salas-de-videoconferencia-unam/>

- Conexión de videoconferencia multipunto H.323/SIP (IP)
- Conexión de videoconferencia multipunto para móviles (WebRTC)
- Sala de videoconferencia.

La UNAM tiene presencia en la mayor parte del territorio de México³¹, y cuenta 12 campus en otros países³², por lo anterior, la inversión en equipo, sistemas, y capacitación de videoconferencia del VNOC de la DGTIC, ha representado ahorros significativos para las finanzas de la UNAM en la realización de las actividades académicas a distancia, principalmente en el pago de pasajes y/o viáticos al campus de Ciudad Universitaria en la Ciudad de México. Como ejemplo, si se realizan reuniones presenciales en el Campus Central con al menos 15 de los directores de las dependencias más alejadas del centro se erogan un promedio de \$5,131.00 USD entre pasajes aéreos y viáticos de hospedaje y alimentación para las 15 personas. Al año \$61,578.00 USD³³ si se realiza al menos una reunión mensual de directores con el Rector.

En los casi 22 años que tiene la RVUNAM, no se había planteado un análisis de la importancia de esta infraestructura en caso de algún problema que no permitiera su operación de manera cotidiana. En primer lugar, la DGTIC cuenta con un manual para continuidad de operaciones en las que se establecen los procedimientos y acciones aplicables por el personal de la DGTIC en caso de contingencias, periodos vacacionales o cualquier otra circunstancia que implique el acceso limitado o restringido a las instalaciones de la dependencia para la realización de las operaciones, así como en situaciones de siniestro en el que no se pueda tener acceso a procesos, servicios críticos para minimizar el impacto que los accesos restringidos pudieran causar y garantizar de esta forma la seguridad física de equipos, recursos, instalaciones y patrimonio universitario a cargo de la DGTIC. El Centro de Operaciones de Videoconferencia no tenía un análisis o un plan de continuidad en caso de desastres. No se tenían establecidos procedimientos de acción específicos a la infraestructura de videoconferencia, ni se había cuantificado el impacto operativo y económico para la realización de las reuniones de trabajo con las sedes que están fuera de la Ciudad de México.

Con base en Bernuy [4], existen diversos estándares para el desarrollo de planes de seguridad informática, algunos de ellos son ITIL (*Information Technology Infrastructure Library*), que es un conjunto de buenas prácticas para la gestión de los servicios de Tecnologías de la Información (TI), COBIT (*Control Objectives for Information and related Technology*) que es un marco de referencia para la aplicación de buenas prácticas en la administración de (TI), o la Norma ISO 17799 que es una guía de recomendaciones de buenas prácticas para la gestión de la seguridad de la información. Para la realización de este trabajo se revisaron algunas prácticas relevantes de estos estándares y normas como vistas en un curso de seguridad informática [5] que el autor asistió en la Maestría de Informática Administrativa impartida en la UNAM³⁴.

El alcance que se tiene para este plan es para la infraestructura física y lógica que permite la operación de los servicios de videoconferencia de la UNAM.

³¹ La UNAM en la República Mexicana <https://www.unam.mx/acerca-de-la-unam/mapas-unam>

³² Presencia de la UNAM en el mundo <https://www.unam.mx/acerca-de-la-unam/mapas-unam>

³³ Estimación propia con costos promedio de aerolíneas, hospedaje y alimentación en México.

³⁴ Maestría en Informática Administrativa, http://posgrado.fca.unam.mx/maestria_informatica.php

1.4 Objetivo del Plan de Continuidad del VNOC

El objetivo del Plan de Continuidad es mantener los servicios del Centro de Operaciones de Videoconferencia mediante el conocimiento de los riesgos potenciales y las acciones que permitan el restablecimiento de los servicios con la menor afectación posible a los usuarios.

La realización de este plan tuvo una duración de realización de 4 meses, ha permitido contar con la identificación de las amenazas y el impacto en la realización de actividades académicas a distancia de la institución.

2. Administración del Plan de Continuidad del VNOC

De manera general, las acciones preventivas que se realizaron para la elaboración del plan son los siguientes: 1) Identificar los usuarios de los servicios del VNOC, 2) Identificar los activos críticos, 3) Identificar y evaluar riesgos a priorizar para su gestión, 4) Tratar los riesgos para minimizar su impacto, 5) elaborar un manual de contingencias y recuperación de desastres disponible en línea e impreso para el personal del VNOC y Dirección de Telecomunicaciones.

Un riesgo es la probabilidad de que algo suceda como resultado de un peligro o amenaza y el impacto que tendrá en los servicios del VNOC. El riesgo surge de la incertidumbre y este se mide en términos de la probabilidad de que suceda y las consecuencias si sucede.

La realización de las siguientes etapas requirió del involucramiento de las áreas operativas, es decir, los Departamentos de Redes y Centro de Datos, así como de las empresas que tienen contratos de mantenimiento celebrado con la institución. Como el problema era coincidir con las agendas de todos para una reunión general, se enviaron correos electrónicos a los responsables una vez que se tenía identificado las actividades que se realizarían. Posteriormente, se hicieron reuniones con cada uno de los responsables para identificar los procedimientos que se tenían que hacer para cada uno de los activos críticos.

2.1 Perfil de usuarios del VNOC

Los servicios de videoconferencia de la DGTIC son operados por el personal del Centro de Operaciones de Videoconferencia con el apoyo de diferentes áreas internas y externas de la DGTIC. La colaboración con áreas internas permite la administración de las configuraciones para los distintos equipos y servicios de red de los que se apoya el servicio de videoconferencia que se brinda a la comunidad universitaria. Con los proveedores externos se tienen contratos de mantenimiento preventivo y correctivo para la prestación de diversos servicios. Y finalmente, los usuarios de los servicios de videoconferencia. En la siguiente tabla se presentan algunos de los diferentes usuarios de la infraestructura lógica y física identificados:

Tabla 1. Perfil de usuarios de los activos físicos del VNOC (Sólo se muestran 5).

Usuarios	Ubicación	Actividades en los activos	Equipos o servicios disponibles	Permisos
Jefe del departamento de redes	Departamento de Operación de la RedUNAM	Administración de switches, routers Direccionamiento IP	Switch capa 3	Leer Escribir Borrar Modificar Borrar
Jefe del departamento del Centro de Datos	Centro de Datos de la UNAM	Administración de VMWare Administración de dir IP de VMWare	Servidor Cisco UCS Máquinas virtuales	Crear Leer Escribir Modificar Borrar
Jefe de Departamento de Comunicaciones Audiovisuales	Centro de Operaciones de Videoconferencia	Administración de servicios de videoconferencia	MCU Tandberg MCU Pexip Codec de videoconferencia Sitio Web VNOC	Crear Leer Escribir Modificar Borrar
Ing. de operaciones de videoconferencia	Centro de Operaciones de Videoconferencia	Operación de MCU Operación de Codec Operación de página web	MCU Tandberg MCU Pexip Sitio Web del VNOC Codec de Videoconferencia	Crear Leer Escribir Modificar Borrar
Proveedor del contrato de mantenimiento del MCU Tandberg y codéc de videoconferencia	Servicio al cliente de la empresa proveedora	Atención a reportes realizados por el VNOC	MCU Tandberg Códéc Cisco SX80	Actualización de las configuraciones Mantenimiento preventivo y correctivo

Una vez identificados los usuarios de los servicios de videoconferencia el siguiente paso es identificar los activos críticos. De acuerdo con la ISO, los activos, son ítems, objetos o entidades que tienen valor real o potencial para la organización. Un activo crítico puede referirse a los activos necesarios para proporcionar servicios a los clientes críticos. Estos activos pueden ser críticos desde el punto de vista de la seguridad, del ambiente, o del desempeño [6].

2.2 Activos críticos de software del VNOC.

Las aplicaciones que son utilizadas por el Centro de Operaciones para brindar los servicios de videoconferencia son aquellos que permiten la administración y configuración de sistemas y equipos; así como brindar servicios de información,

realización de solicitudes por parte de los usuarios finales. En la siguiente tabla se presentan los principales utilizados para este servicio.

Tabla 2. Activos críticos de software del VNOC (Sólo se muestran 8).

ID	Activo	Descripción
1	Sistema operativo del MCU Tandberg MPS800	Software que permite la operación y administración del hardware de la unidad multipunto.
2	Consola de administración del MCU Tandberg MPS800	Consola de acceso vía web para programar y monitorear las conferencias multipunto de este equipo.
3	Sistema operativo del códec Cisco SX80	Software que permite la operación y administración del equipo de videoconferencia Cisco SX80.
4	Correo electrónico	Aplicación que permite la comunicación interna y externa con los usuarios y personal de diversas áreas de la institución.
5	Servicio de internet	Servicio que permite las comunicaciones digitales entre los usuarios del servicio de videoconferencia.
6	Sistema VMWare del servicio MCU PEXIP	Software que permite virtualizar un servidor en varias máquinas para implementar un nodo de videoconferencia multipunto con la aplicación PEXIP.
7	Sistema operativo del MCU PEXIP	Software que permite la operación y administración del hardware del sistema multipunto PEXIP.
8	Consola de administración del MCU PEXIP	Consola de acceso vía web para programar y monitorear las conferencias multipunto en el sistema PEXIP..

2.3 Activos críticos de hardware del VNOC

En el Centro de Operaciones de Videoconferencia de la DGTIC-UNAM los activos de hardware son recursos esenciales para brindar los servicios de conectividad multipunto.

Tabla 3. Activos críticos de hardware del VNOC (sólo se muestran 5):

ID	Activo	Descripción
1	Codec de videoconferencia Cisco SX80	Equipo instalado en la sala de videoconferencia para brindar el servicio de comunicaciones audiovisuales grupales.
2	MCU Tandberg MPS800	Equipo de infraestructura central que permite la comunicación de 2 o más sistemas de videoconferencia. Instalado en el sitio de telecomunicaciones de la Zona Cultural de Ciudad Universitaria.
3	Servidor Cisco UCS	Equipo instalado en el centro de datos de la DGTIC que contiene instalado el nodo de administración y nodo de videoconferencia del sistema multipunto PEXIP.
4	Switch Cisco 24 puertos	Equipo instalado rack de videoconferencia que permite la conectividad de red local y acceso a internet de los equipos operados por el personal del departamento.
5	PC de escritorio	Equipo utilizado para las actividades de control y monitoreo de los servicios del departamento.

3. Análisis de riesgos de activos del VNOC

El riesgo informático se define como una dificultad que interviene en el cumplimiento de una meta o una amenaza para la pérdida de la información.

El análisis de riesgos tiene por objetivo estudiar amenazas que tienen probabilidad de afectar los servicios de los activos físicos y lógicos del Centro de Operaciones de Videoconferencia de la DGTIC-UNAM como unidades de control multipunto de videoconferencia, servidores de VMWare para la virtualización de servicios multipunto, equipos terminales de video, servidores web, y sistemas de información.

Para realizar el análisis de riesgo se elaboraron matrices que ayudan a calificar la probabilidad de un riesgo y el impacto que afectaría el funcionamiento de los servicios del Centro de Operaciones, así como a la integridad de la información resguardada en sus sistemas.

Las matrices se clasifican de la siguiente manera:

- Activos físicos: todos los equipos, servidores, computadoras y elementos

de red operados por el Centro de Operaciones de Videoconferencia.

- Activos lógicos: Sistemas de información, páginas web, sistemas operativos y servicios de red que son necesarios para brindar los servicios del Centro de Operaciones de Videoconferencia.
- Usuarios: Todas las personas que tienen acceso a los equipos, servicios e información del Centro de Operaciones de Videoconferencia.
- En este análisis de riesgos se identificaron los activos del Centro de Operaciones, y sus amenazas. Se utilizaron las siguientes variables y valores:
- Riesgo: Escala de 1 a 4 a mayor valor determina inmediatez en la necesidad de atención.
- Probabilidad: Escala de 1 a 5 mide la probabilidad de incidencia.
- Consecuencias: Escala de 1 a 5 mide el impacto del daño probable.

Tabla 7. Valores utilizados en el análisis de riesgo.

Riesgos		Probabilidad		Consecuencias	
Muy alto	4	Certeza	5	Catastróficos	5
Alto	3	Probablemente	4	Mayores	4
moderado	2	Moderado	3	Moderados	3
Bajo	1	Improbable	2	Menores	2
		Raro	1	Insignificantes	1

Estas variables se multiplicarán para determinar qué riesgos son prioridad. La escala que se utilizó es la siguiente:

Tabla 7. Escala de valores y colores de riesgos.

Color	Rango de valores	Significado
Muy Alto	60 - 100	Requiere atención inmediata
Alto	50 - 59	Debe atenderse a la brevedad posible
Medio	30 - 49	Puede postergarse, pero debe atenderse
Bajo	10 - 29	Problema secundario, puede esperar
Muy Bajo	0 - 10	Altamente probable

En la siguiente tabla, se presenta el análisis de riesgos en los activos lógicos y físicos con los riesgos más altos detectados en el VNOC, que como se comentó, los números asignados en cada columna, es una estimación que realizamos los responsables de los sistemas en base a la experiencia que se ha tenido con otros sistemas similares, sus incidencias y la probabilidad que suceda el problema descrito:

Tabla 8. Análisis de riesgos de activos lógicos y físicos del VNOC UNAM (Solo se muestran un par de los elementos analizados).

Activos	Riesgo	Descripción	Impacto	Niv. de Riesgo	Niv. de Probabilidad	Niv. de Consecuencias	Total
Sistema VMware del servicio MCU PEXIP	Falla en la carga del nodo multipunto	Falla en la carga de servicios multipunto	Error en la carga y conexión de sistemas multipunto.	4	3	3	36
Consola de administración del MCU PEXIP	Falla en los respaldos de información de configuración de consola MCU PEXIP	Error en el respaldo de la información de la consola MCU PEXIP	Pérdida de configuraciones	5	2	5	50

La forma en que se obtiene el resultado final en la última columna Total es una multiplicación de cada uno de los valores de la columna. En el caso del activo de la consola de administración la siguiente operación indica el resultado:

$$(\text{Riesgo} \times \text{probabilidad} \times \text{nivel de consecuencias}) = 5 \times 2 \times 5 = 50 . \quad (1)$$

4. Manual de contingencias y recuperación de desastres

Un plan de contingencias es un conjunto de procedimientos alternativos a la operatividad normal de la organización. La finalidad es la de permitir el funcionamiento de esta, aun cuando algunas de sus funciones dejen de hacerlo en caso de algún incidente interno o externo.

4.1 Objetivo del manual de contingencias y recuperación de desastres

El objetivo general de este manual de contingencias es brindar una guía de acciones a seguir para el restablecimiento de los servicios del Centro de Operaciones de Videoconferencia en caso de desastre.

Los objetivos específicos del plan de contingencia para el VNOC son:

- Establecer un procedimiento formal y por escrito que indique las acciones a seguir frente a los riesgos identificados en los activos críticos.
- Optimizar el uso de recursos humanos y materiales.
- Contar con un control adecuado para cumplir las normas y procedimientos establecidos

Para lograr lo anterior, de acuerdo con el análisis de riesgos, se plantearon estrategias y acciones a seguir sobre los activos críticos que se identificaron con mayores riesgos. En algunos casos se recomiendan acciones preventivas y en otros casos se proponen acciones correctivas. Con este manual, se brinda al personal de TI y del VNOC ejecutar un conjunto de acciones básicas de respuesta que se deberían de tomar para afrontar de manera oportuna, adecuada y efectiva, ante la eventualidad de incidentes, accidentes y/o estados de emergencias que pudieran ocurrir tanto a las instalaciones o fuera de ella.

4.2 Estrategias para activos críticos de software

Tabla 9. Manual de contingencias de activos lógicos del VNOC UNAM (Solo se muestran uno de los elementos analizados).

Activo		
Sistema VMware del servicio PEXIP MCU	Riesgo	Falla en la carga del nodo multipunto
	Descripción	Falla en la inicialización del sistema VMWARE en el nodo de administración.
	Impacto	Error en la carga y conexión de sistemas multipunto.
	Estrategia	Realizar un plan de inicio secundario en caso de falla.
	Tipo	Preventiva
	Acciones a seguir	-Revisar el estado de los servicios con el Centro de Datos -En caso de ser requerido, levantar el reporte con la empresa responsable del mantenimiento de los servicios de VMWARE -Documentar las actividades realizadas.
	Responsable	Jefe del Depto de Comunicaciones Audiovisuales, Ingeniero de Operaciones, Jefe del Centro de Datos

4.3 Estrategias para activos críticos de hardware

Tabla 10. Manual de contingencias para activos físicos (Sólo se muestra un activo analizado).

Activo		
MCU Tandberg MPS800	Riesgo	No se tiene acceso vía web a la consola de administración
	Descripción	Cuando se requiere operar el equipo no es posible abrir la página web de la consola de administración
	Impacto	Falla completa del MCU, no es posible dar servicios de conectividad por videoconferencia multipunto
	Estrategia	Contar con el contrato de mantenimiento actualizado con el proveedor del servicio
	Tipo	Preventiva
	Acciones a seguir	Verificar con el departamento de redes que no se tengan restricciones de acceso. Si es que esto es negativo, levantar un ticket de servicio al número de teléfono del proveedor del servicio de mantenimiento preventivo y correctivo del equipo. Verificar con el proveedor que se realicen las acciones necesarias para reestablecer el equipo. Si está funcionando correctamente el acceso web a la consola de administración cerrar el ticket.
	Responsable	Jefe del Departamento VNOC y proveedor del contrato de mantenimiento del MCU.

Resultados y conclusiones

El Plan de Continuidad del VNOC ha permitido identificar cuáles son los activos críticos de esta área para poder establecer su nivel de importancia en la atención de posibles incidentes, así como tener identificados a las personas y las áreas que son responsables de la configuración y operación de la infraestructura tecnológica.

Los contratos de mantenimiento de los equipos de videoconferencia han sido en algunas ocasiones un problema financiero para las Instituciones de Educación Superior y principalmente en la UNAM. El costo de no tenerlas es mayor en caso de algún problema o contingencia del equipo. En el caso de la MCU Tandberg que soporta 80 puertos H.323, su costo en 2006 fue de aproximadamente \$120,000.00 USD, con los 12 años de uso que se le ha dado al equipo se estima que tiene un costo de \$10,000.00 USD por año. El costo de un contrato de mantenimiento preventivo y correctivo anual es de \$34,274.00 USD y que por los 12 años da un total de \$411,299.00 USD. Sumados, el valor del MCU en 12 años más las pólizas de mantenimiento da un total de \$531,299.00 USD. Si se estima que al menos las reuniones de directores se hacen una vez al mes, y que al año da un costo aproximado de \$61,578.00 USD, en 12 años se sumarían \$738,936.00 USD por conceptos de pasajes aéreos y viáticos de hospedaje y alimentación de 15 directores. Entonces, la diferencia entre el costo del MCU y el costo de las reuniones presenciales es de \$207,637.00 USD por 12 años, y \$17,303.00 USD anual, que representaría el ahorro en la realización de reuniones virtuales versus las reuniones presenciales. Es por esto que uno de los pilares del plan de continuidad son los contratos de mantenimiento que permiten garantizar la operación de un equipo que brinda conectividad multipunto a la institución.

Adicionalmente, el Plan de Continuidad es sumamente útil para la capacitación del personal que se pueda incorporar al VNOC debido a que, en conjunto con el manual de organización y operación, permite identificar actividades y procedimientos en caso de contingencias o desastres en la infraestructura de videoconferencia.

Con la revisión de la documentación relativa a seguridad lógica en los servicios de Internet del VNOC, se implementó el protocolo HTTPS (*Hyper Text Transfer Protocol*) en el sitio web del Departamento que ahora brinda comunicaciones cifradas hacia los usuarios [7]. Con el apoyo del Departamento de Firma Electrónica Avanzada de la UNAM [8] y el Centro de Datos, se realizó la solicitud de un Certificado Digital (CSR por sus siglas en inglés) y un certificado SSL. Mediante esta iniciativa, se protege el sitio web con el certificado SSL y se mantienen comunicaciones seguras en la transferencia de hipertexto al utilizar el protocolo HTTPS.

Es importante comentar que el Plan de Continuidad no es infalible ya que se pueden suceder nuevas incidencias en la seguridad física o lógica a los equipos centrales de videoconferencia que no se tenían detectados. Sin embargo, es un punto de referencia para brindar acciones de seguridad informática, así como para continuar actualizando cada una de sus secciones o agregar las que seas necesarias. De igual forma, se podrían quitar las que ya no son relevantes o si los equipos se han retirado de la operación por obsolescencia.

Con el Manual de Contingencias y Recuperación de Desastres, el personal de TI y del VNOC no tiene que esperar a que se encuentre el Jefe del Departamento o responsable del área porque tiene establecidas las tareas necesarias para reestablecer algún servicio en conjunto con los departamentos que apoyen los servicios del VNOC,

principalmente el Departamento de Redes y el Centro de Datos.

El hecho de contar con un Plan de Seguridad en una organización no es solo tenerlo en el estante o librero, también es importante que se comunique a todo el personal involucrado en TI de manera escrita y visual. Es importante que el personal entienda cómo minimizar las vulnerabilidades de seguridad. En conclusión, todo plan de seguridad requiere de revisión y actualizaciones, debido a que los riesgos en la seguridad física y lógica cambian constantemente.

Agradecimientos

El autor desea expresar su agradecimiento a las autoridades y colegas de la UNAM, así como las personalidades de las siguientes instituciones por las facilidades otorgadas y su participación en el desarrollo de este proyecto:

- Dirección General de Cómputo y de Tecnologías de Información y Comunicación de la Universidad Nacional Autónoma de México
- Dr. Felipe Bracho Carpizo – Director
Dirección de Telecomunicaciones
 - o M. en C. Ma. de Lourdes Velázquez Pastrana
 - o Ing. Roberto Rodríguez Hernández
- Dirección de Sistemas y Servicios Institucionales
 - o Act. Fabián Romo Zamudio
 - o Ing. Pedro Bautista
- Corporación Universitaria para el Desarrollo de Internet
 - o Lic. Carlos Casasús
 - o Lic. Martha Ávila
- Posgrado de la Facultad de Contaduría y Administración
 - o Mtro. Rene Montesano
 - o Lic. Rosalba Núñez
 - o Lic. José Rodrigo Hernández
 - o Lic. Rosario Rivas

Referencias

1. Dirección General de Cómputo y de Tecnologías de la Información y Comunicaciones, <https://www.tic.unam.mx>
2. Wainhouse Research. Understanding and Deploying a VNOC., (2008): <https://slidex.tips/download/understanding-and-deploying-a-conferencing-vnoc>
3. Departamento de Comunicaciones Audiovisuales <https://vnoc.unam.mx>
4. Bernuy, J, Tesis: Implementación de seguridad de la información mediante ISO-17799 plataforma BSD. Programa de Maestría y Doctorado en Ingeniería, UNAM. pp 6 a 14 (2008).
5. Montesano, R, "Importancia del análisis de riesgo en la planeación de la seguridad informática". Seguridad en Informática [diapositivas de Power Point], Maestría en Informática Administrativa, Facultad de Contaduría y Administración, UNAM.

Diapositiva 5 a 16, (2008)

6. ISO, Online Browsing Platform (OBP), Gestión de activos – Aspectos generales, principios y terminología, <https://www.iso.org/obp/ui#iso:std:iso:55000:ed-1:v2:es>
7. El cifrado web (SSL/TLS) <https://revista.seguridad.unam.mx/numero-10/el-cifrado-web-ssltls>
8. Departamento de Firma Electrónica Avanzada, <https://www.fea.unam.mx/>

Combate à Fraudes em Documentos Acadêmicos: Registro, Autenticação e Preservação de Diplomas Digitais Usando *Blockchain*

Rostand Costa¹, Daniel Faustino² e Guido Lemos¹

¹*Laboratório de Aplicações de Vídeo Digital - LAVID
Universidade Federal da Paraíba - UFPB*

²*Centro de Ciências Exatas e Naturais - CCEN
Universidade Federal Rural do Semiárido - UFERSA*

Linha Temática: Logrando universidades digitales seguras.

Abstract: The purpose of this paper is to investigate the potential use of the blockchain technology combined with active distributed repositories to create a platform, scalable and agnostic³⁵, specialized in the authentication and preservation of digital documents. As a proof of concept of the proposed platform, was performed the construction of a public service for digital registration and verification of the authenticity of academic documents. The prototype of the service offers an interface for educational institutions to register official documents, such as diplomas and certificates, using blockchain and an interface so that users can verify the authenticity of a document through its registration number in a DLT. The documents registered in the service are automatically inserted in the long-term digital preservation repository.

³⁵ Agnóstica, no contexto deste documento, significa independência do formato de representação e do tipo de conteúdo armazenado no objeto digital a ser registrado e preservado.

1. Introdução

No esteio da metamorfose digital que a sociedade como um todo vem passando na última década, o ecossistema universitário também tem sido palco e alvo das pressões inovadoras que esta onda traz. Aplicada aos processos e métodos, a transformação digital pode tornar as empresas mais eficientes e no segmento de educação não é diferente. Este movimento evolutivo alimenta e é alimentado por vários atores que interagem, dentre eles a comunidade, as empresas e o governo.

Neste sentido, nos últimos anos, o Ministério da Educação do Brasil tem publicado uma série de Decretos e Portarias, com novos requisitos tecnológicos (e, eventualmente, também procedurais) que precisam ser cumpridos pelas Instituições de Ensino Superior – IES e que visam promover a padronização da gestão de documentos, aprimorar a disponibilidade, a rastreabilidade e integridade do acervo acadêmico.

Este artigo traz os resultados preliminares do Grupo de Trabalho GT-RAP³⁶, um projeto de P D financiado pela Rede Nacional de Pesquisa e Ensino (RNP) e executado pelo Laboratório de Aplicações de Vídeo Digital (LAVID). O objetivo do GT-RAP é investigar o potencial do uso combinado das tecnologias de *blockchain*, certificação digital e preservação digital para a criação de uma plataforma, escalável e agnóstica, especializada no registro, autenticação e preservação de documentos digitais relevantes (DDRs).

Como prova de conceito da plataforma proposta, foi desenvolvido e implementado um serviço público de registro digital e verificação da autenticidade de documentos acadêmicos, como diplomas e certificados, para combater as diversas classes de fraudes associadas a esse tipo de documento. A escolha deste contexto deveu-se à grande ocorrência de fraudes ao redor do mundo, motivados principalmente pela facilidade de falsificação e pelos ganhos financeiros envolvidos. No Brasil, não é diferente e a evidência da escala do problema levou o Ministério da Educação a promulgar uma legislação que institui recentemente uma série de medidas que devem ser tomadas pelas universidades para enfrentar o problema.

Em fevereiro de 2019, a primeira emissão de diplomas acadêmicos digitais foi realizada no Brasil usando registro em *blockchain* através da Plataforma RAP. A UFPB, Universidade Federal do Nordeste do Brasil, emitiu os diplomas dos egressos de 2018 dos cursos do Centro de Informática. Várias outras instituições já expressaram o desejo de usar a Plataforma RAP, que é aberta, extensível e suporta múltiplos emissores e várias classes de documentos.

Como trabalho futuro, gostaríamos de propor um modelo de autenticação federada para DDRs, especialmente diplomas, que permita a colaboração internacional na luta contra a fraude. Este artigo tem como objetivo divulgar a iniciativa e disponibilizá-la para a comunidade.

O restante do documento está organizado como segue. A Seção 2 traz a contextualização e motivação do projeto, incluindo uma breve discussão sobre ocorrências de fraudes associadas com diplomas acadêmicos, como a gestão de tais documentos ocorre e como anda o processo de adoção de diplomas digitais. Na Seção 3 é feita a apresentação da plataforma proposta e também é discutida a estratégia de modelagem adotada, de forma a permitir conciliar a evolução tecnológica sugerida com

³⁶ <http://gt-rap.lavid.ufpb.br>

as transformações culturais e regulatórias necessárias. Na Seção 4 são discutidos alguns aspectos chaves da tecnologias que foram combinadas para construir a plataforma proposta. Na Seção 5 é realizado o detalhamento de como o protótipo da plataforma foi projetado e construído, descritos resultados preliminares de integração com duas IES de grande porte e também a geração pioneira de diplomas digitais no Brasil usando *blockchain*. A Seção 6 traz as nossas considerações finais e uma breve análise das lições aprendidas.

2. Problemática e Contexto

2.1 Necessidade de Proteção de Documentos Acadêmicos

Os últimos dados³⁷ do Censo da Educação Superior realizado anualmente pelo INEP indicam que o Brasil teve em 2016 um total de um milhão e cem mil concluintes e um total de quase três milhões de ingressantes nas instituições públicas e privadas de ensino superior. Com isso, em 2016, o total de estudantes matriculados em cursos de ensino superior no país ultrapassa pela primeira vez os 8 milhões de alunos.

Os números³⁸ da CAPES sobre os programas de pós-graduação no Brasil indicam que em 2016 haviam 325.320 estudantes matriculados em cursos de pós-graduação *stricto sensu* no país, variando entre mestrado profissionalizante, mestrado acadêmico e doutorado. Neste mesmo ano foram titulados no Brasil 20.630 doutores e 59.349 mestres.

Em relação aos cursos de pós-graduação *lato sensu*, não há dados atualizados sobre a quantidade de cursos em funcionamento ou sobre a quantidade de alunos matriculados e titulados. São números bastante variáveis dada a natureza mais dinâmica desses cursos, sempre associados a necessidades do mercado, e também devido as baixas exigências regulamentares para sua abertura. Porém, é possível estimar, de forma bastante conservadora, que a pós-graduação *lato sensu* no Brasil seja, pelo menos, 10 vezes maior que a pós-graduação *stricto sensu*, o que nos levaria a diplomação de pelo menos 600.000 especialistas ao ano.

Com isso, podemos chegar a conclusão que no Brasil, apenas no ensino superior, são emitidos anualmente cerca de 1,8 milhão de diplomas. Se somarmos a este número a quantidade de diplomas emitidos no exterior e revalidados no país, é possível atingirmos a marca de 2 milhões de diplomas emitidos anualmente.

Apesar da emissão de tais diplomas ser controlada, exigindo-se o registro do documento por uma Universidade junto ao Ministério da Educação, as dimensões continentais de nosso país, aliadas à falta de suporte tecnológico, fazem com que a tarefa de verificar a autenticidade de um diploma emitido ou revalidado no país seja feita de forma ineficiente.

Um exemplo dessa dificuldade envolveu o INEP, um órgão do próprio Ministério da Educação, responsável por diversas atividades relacionadas a avaliação do ensino no

³⁷ <http://portali.inep.gov.br>

³⁸ <http://www.capes.gov.br>

país. Em 2011 o INEP foi obrigado a cancelar as avaliações de 4 cursos de graduação na área de Direito pelo fato de que um dos avaliadores *ad hoc* participantes do **Banco de Avaliadores do Sistema Nacional de Educação Superior** (BASIS) teria apresentado diplomas falsos de mestrado e doutorado [3]. Tal fraude só foi descoberta por conta de uma denúncia anônima. Não fosse isso, possivelmente este avaliador teria continuado a representar o Ministério da Educação na avaliação do ensino superior no país.

Casos de uso de diplomas falsos para ingresso no ensino superior e no serviço público também não são raros. O Ministério Público do Paraná investigou o uso de mais de 500 diplomas falsos na cidade de Maringá que foram utilizados para ingressos em universidades e aprovação em concursos públicos [4].

Há quadrilhas especializadas em vender diplomas falsos em nome de Instituições de Ensino Superior [7]. No estado do Espírito Santo, mais de 100 professores estão sendo processados por usarem diplomas falsos [5]. As fraudes também acontecem no processo de revalidação de diplomas estrangeiros. Um caso emblemático, foi a quadrilha que revalidava de forma fraudulenta diplomas de medicina no estado do Mato Grosso [6]. Além disso, há inúmeros relatos de casos de utilização de diplomas falsos para ludibriar clientes e, o que é mais grave, pacientes, no caso dos diplomas relacionados às áreas de saúde.

O processo de investigação, descoberta e tomada de providência em relação à falsificação de diplomas também envolve desafios específicos. Como os mecanismos de controles e verificação nem sempre são automatizados e o número de casos têm crescido nos últimos anos, é complicado para a justiça detectar e tomar as devidas providências frente aos prejuízos causados por essas quadrilhas e pessoas de má fé. A Secretaria de Estado da Educação do Espírito Santo (SEDU-ES), por exemplo, relata que as investigações referentes aos processos administrativos para apurar o uso de diplomas falsos nas escolas do estado podem durar até 180 dias, com possibilidade de prorrogação [5].

2.1.1 Uso de Diplomas Digitais

O uso de diplomas digitais apresenta uma série de vantagens quando comparado ao diploma tradicional: i) a eliminação, pelo emissor, do custo de impressão da versão em papel (normalmente especial e de alto custo); ii) a replicação e distribuição ilimitada e gratuita do documento pelo portador; e iii) a possibilidade de adoção, pelo destinatário, de mecanismos automatizados de verificação da autenticidade do documento. Tais características emprestam mais agilidade e segurança para que portadores de diplomas e interessados possam compartilhar e verificar a legitimidade de tais documentos de forma mais eficiente. Além disso, é possível padronizar e formalizar o processo de emissão e validação dos diplomas, dificultando tentativas de fraudes.

Embora o uso de diplomas digitais desponte como um caminho quase que natural para a gestão de documentos dessa natureza, ainda não existe uma legislação que trate diretamente do tema. A falta de uma regulação específica, assim como aspectos culturais, atrasavam o avanço do processo de digitalização desses documentos e o seu uso em larga escala. Este cenário está mudando no Brasil.

O MEC, em um movimento coordenado, começou a direcionar em suas portarias um movimento em direção à digitalização. O Decreto 9.235, por exemplo, publicado em 15/12/2017, exige que as Instituições de Ensino Superior (IES) do Brasil direcionem seu acervo acadêmico para o meio digital. A Portaria 315/18, de 04/04/18, tem como objetivo garantir que todas as informações referentes à vida acadêmica dos estudantes sejam gerenciadas e armazenadas digitalmente.

Com relação à emissão do diploma digital, de forma mais específica, o MEC lançou as portarias 330/18, em 05/04/2018, e 554/19, em 12/03/2019. Baseadas na utilização de Certificação Digital ICP-Brasil, as portarias regulam a digitalização dos diplomas acadêmicos de forma a garantir autenticidade, rastreabilidade e validade jurídica. As IES brasileiras terão um prazo de 24 meses para implementar o diploma digital a partir da publicação da Portaria No. 554/19.

2.2 Desafios da Preservação Digital de Longo Termo

Na mesma proporção em que uma parte considerável dos artefatos relacionados à diversas atividades humanas está sendo criada em formatos digitais, é esperado que as práticas de preservação para essas informações também devam ser baseadas em técnicas e tecnologias adequadas e igualmente digitais.

Quando comparada com a preservação de coleções físicas, a preservação de conteúdo digital traz, em si, uma associação, quase paradoxal, de um grande potencial de risco e um grande potencial de proteção [16]. O potencial de risco é representado pela efemeridade do armazenamento digital que pode ser irremediavelmente perdido por causa de falhas técnicas ou humanas com muito mais facilidade e rapidez do que no caso de representações físicas de conteúdo. O potencial de proteção, por sua vez, é ancorado no fato de que coleções digitais podem ser indefinidamente reproduzidas e armazenadas com total fidelidade e integridade.

A perfeita continuidade de coleções digitais depende, em grande parte, de se buscar um equilíbrio da aplicação de medidas que aproveitem ao máximo o potencial de proteção ao ponto de neutralizar o seu inerente potencial de risco. Entretanto, o desafio pode representar muito mais um problema social e institucional do que uma questão meramente técnica, pois, em particular, para a preservação digital no meio acadêmico, depende-se de instituições que passam por mudanças de direção, missão, administração e fontes de financiamento [14].

Além disso, a preservação digital envolve desafios essencialmente diferentes dos encontrados na conservação de conteúdo em suportes mais tradicionais. De um ponto de vista mais tradicional, o ato de preservar traduz-se no ato de manter imutável e intacto. No ambiente digital, entretanto, a ação de preservar também pode se referir a mudar, recriar e renovar. Onde renovar pode significar mudar formatos, atualizar mídias e/ou substituir *hardware* e *software*. Em suma: se, por um lado, queremos manter o conteúdo exatamente como foi criado, intacto; por outro lado, queremos continuar acessando-o em plataformas modernas. Este é o *Paradoxo da Preservação Digital*, conforme descrito por Sayão [14].

Um número crescente de organizações de memória cultural (incluindo as reunidas na iniciativa *MetaArchive* [15]) aposta que os esforços mais eficazes de preservação digital ocorrem na prática através de alguma estratégia para manter múltiplas cópias de

conteúdo digital em locais distribuídos seguros[13, 10]. Na era digital, esta estratégia requer investimentos numa matriz distribuída de servidores, capazes de armazenar coleções digitais em uma metodologia pré-coordenada.

Entretanto, é pouco provável que uma única organização educacional tenha a capacidade de operar de forma adequada vários servidores distribuídos geograficamente. Neste sentido, a colaboração entre instituições é essencial, e tal colaboração exige investimentos técnicos e organizacionais [1, 2]. Não é o caso de contar apenas com uma solução tecnológica adequada, mas também precisam ser estabelecidos acordos interinstitucionais robustos de longo prazo, ou não haverá compromisso suficiente para uma atuação sintonizada ao longo do tempo.

3. Descrição da Solução Implementada

3.1 Serviço de Autenticação e Preservação Digital de Documentos Acadêmicos

É baseado neste cenário que propomos a construção de uma plataforma, escalável e agnóstica, para armazenamento e verificação digital da autenticidade de documentos digitais baseada no uso combinado das tecnologias de *blockchain*, certificação digital e preservação digital.

A estratégia utilizada como metodologia para o desenvolvimento do projeto previu a divisão do esforço em três grupos de atividades, os quais estavam relacionados ao levantamento do estado da arte e mapeamento de requisitos para certificação digital de documentos e preservação digital distribuída, definição de uma arquitetura genérica para autenticação e preservação de documentos digitais e a montagem de um protótipo para validação da abordagem proposta. Uma visão geral do contexto do serviço e dos principais atores serviço ilustrado na Figura 1.



Figura 1: Visão Geral do Serviço Proposto

O serviço proposto não interfere no fluxo natural entre os três atores em pauta mas permite que tanto a atuação do emissor quanto do receptor do documento digital assinado seja facilitada, sobretudo na interação para a validação do mesmo. Sempre centrado no uso do documento digital assinado pelos atores, o serviço provê mecanismos para registrar e preservar o documento para o emissor e para autenticá-lo para o receptor, fazendo ainda a guarda de longa duração do mesmo, o que beneficia todos os atores.

No contexto do piloto, focado em diplomas acadêmicos, o emissor do diploma digital é a IES de Origem, o portador do diploma é o aluno egresso da IES de Origem e o papel do receptor é representado por outras IES, órgãos públicos e empresas privadas para as quais o portador do diploma está aplicando para uma posição que exige a apresentação do mesmo.

Acreditamos que com um serviço como o proposto aqui em funcionamento, a autenticidade de documentos (como diplomas ou certificados) emitidos ou revalidados por instituições acadêmicas brasileiras poderá ser mais facilmente verificada, tanto por órgãos públicos quanto por instituições privadas e pessoas físicas.

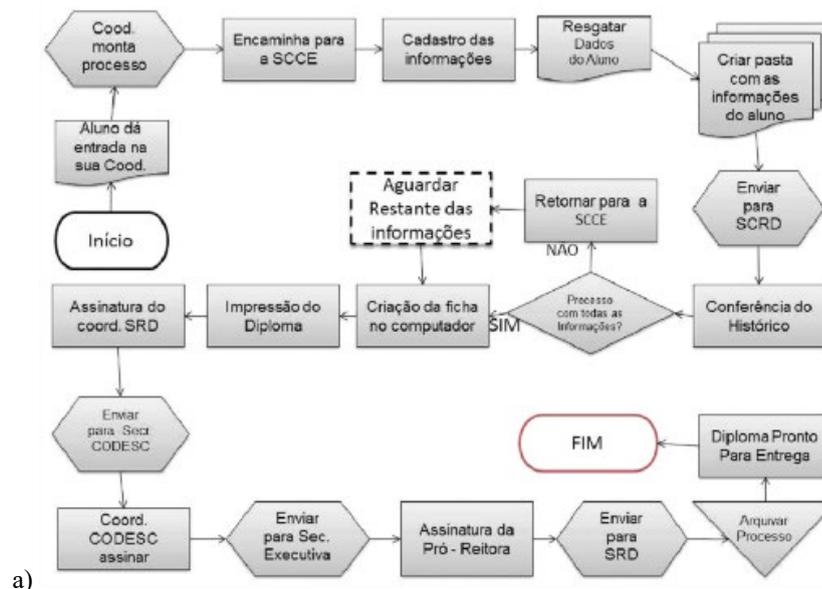
3.2 Estratégia de Modelagem

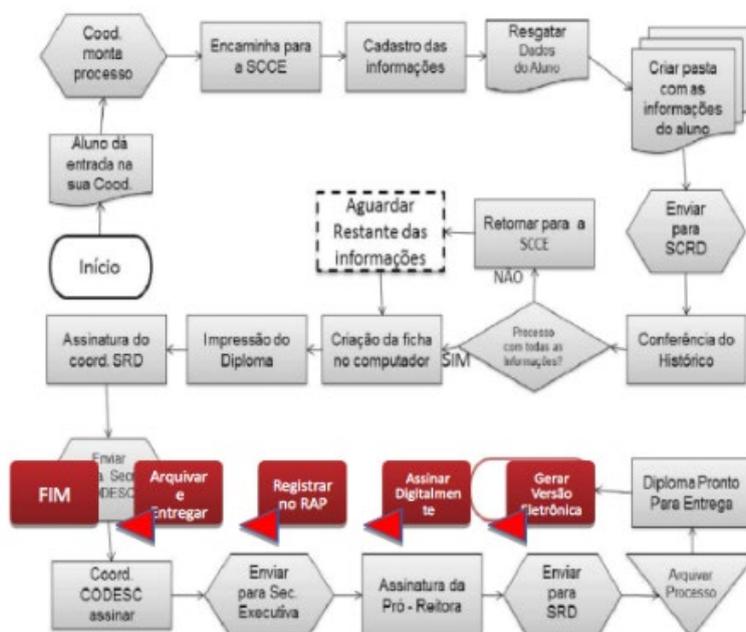
A estratégia de modelagem do protótipo partiu da premissa de garantir uma integração pouco invasiva com os processos e sistemas em uso atualmente para emissão e registro de diplomas acadêmicos nas IES. Considerando a necessidade de uma transição suave para permitir o acultramento progressivo com as novas metáforas e tecnologias envolvidas, considerou-se a possibilidade de uma possível convivência dos dois métodos de emissão de diplomas: tradicional em papel e no formato digital.

Neste sentido, alguns requisitos básicos foram estabelecidos para essa primeira fase do protótipo:

- Tentar minimizar a necessidade de intervenção nos fluxos internos dos setores envolvidos na emissão e registro de diplomas das instituições;
- Considerar que o diploma tradicional, em papel, continuará sendo emitido normalmente e que o diploma digital será uma nova opção para o aluno;
- Identificar e aplicar os mesmos protocolos e níveis de alçada usados para a assinatura tradicional dos diplomas em cada instituição na assinatura digital da versão eletrônica;
- Garantir que todo o controle e autonomia presentes hoje nas IES permaneçam inalteradas para a assinatura digital da versão eletrônica. Não deve existir nenhuma transferência de alçada e responsabilidade para o serviço proposto;
- As operações de registro e autenticação de documentos devem ser sempre lastreadas pela validação das assinaturas digitais dos documentos digitais envolvidos e não apenas baseadas na autenticação de usuários e sessões.

Para ilustrar melhor o contexto, a Figura 2 (a) traz o fluxo operacional atualmente praticado na UFPB para a emissão de diplomas. Como pode ser visto, são feitas três assinaturas no diploma durante o processo de emissão e por atores/setores diferentes.





b)
Figura 2: Fluxo Operacional para Emissão de Diplomas na UFPB (a) e Possível Fluxo de Transição (b)

Para permitir uma conciliação com tais fluxos internos, que variam de instituição para instituição, a integração com o serviço proposto pode iniciar apenas a partir do momento em que a emissão do diploma tradicional em papel é concluída (Figura 2 (b)). Isso garante que todas as verificações do mérito e das exigências do título acadêmico já foram satisfeitas e concluídas.

A partir deste ponto e contando com o desejo e anuência expressa do interessado, o processo adicional de emissão e registro do diploma digital pode ser iniciado pela própria IES. Esta fase possui cinco etapas lógicas bem definidas:

- Geração de uma versão eletrônica do diploma do aluno;
- Assinatura digital da versão eletrônica pelo(s) mesmo(s) responsável(is) formal(is) na instituição que assinaram a versão impressa;
- Submissão da versão assinada para registro na DLT e preservação digital no serviço proposto;
- Recepção e arquivamento do recibo de registro na DLT;
- Entrega ao aluno da versão eletrônica assinada do diploma e também do recibo do registro eletrônico.

Estas cinco etapas podem ser implementadas de diferentes formas em cada instituição. Dependendo da maturidade tecnológica e da disponibilidade de recursos, esses passos podem ser integrados ao sistema de controle acadêmico da instituição ou podem ser tratados em uma aplicação específica, por exemplo.

Para permitir a integração, o protótipo de serviço oferece uma interface programável (API) para que instituições de ensino possam registrar (e também validar) diplomas acadêmicos nas DLTs desejadas. Há também uma interface interativa (Portal) para que

os interessados possam verificar a autenticidade de um documento através do seu número de registro.

4. Aspectos Críticos e Relevantes

A validação da existência ou da posse de documentos formalmente assinados é fundamental em qualquer contexto legal. Normalmente, a certificação tradicional de documentos físicos se baseia em autoridades centrais, notariais ou não, para armazenar e aplicar os registros e mecanismos necessários para tal fim e também lidar com os aspectos e desafios da segurança. Desafios esses que se tornam cada vez mais difíceis à medida que os arquivos envelhecem.

A validação da existência ou da posse de documentos formalmente assinados é fundamental em qualquer contexto legal. Normalmente, a certificação tradicional de documentos físicos se baseia em autoridades centrais, notariais ou não, para armazenar e aplicar os registros e mecanismos necessários para tal fim e também lidar com os aspectos e desafios da segurança. Desafios esses que se tornam cada vez mais difíceis à medida que os arquivos envelhecem.

Entretanto, a materialização e desmaterialização de documentos digitais relevantes (DDRs) bem como o dinamismo e velocidade das relações digitais têm representado uma nova frente para entidades emissoras e/ou certificadoras de documentos formais. Principalmente quando começa a emergir a possibilidade de geração de documentos em papel a partir de documentos digitais relevantes e a geração de documentos digitais relevantes a partir de documentos em papel, demandando a garantia de que os termos estabelecidos no original sejam efetivamente conservados e recebam uma chancela de legitimação, independentemente da sua forma de representação.

De um ponto de vista prático, a certificação de documentos digitais apresenta três dimensões principais [9]: i) *Prova de Propriedade/Autoria* (quem é o detentor/autor do documento), ii) *Prova de Integridade* (o documento está íntegro e exatamente igual à quando foi criado) e iii) *Prova de Existência* (o documento foi criado e legitimado em um dado momento no tempo).

Neste sentido, a tecnologia de livros-razão distribuídos [11] (ou DLTs, do inglês *Distributed Ledgers Technologies*), normalmente baseados em *blockchain*, se apresenta como um modelo alternativo para a certificação de documentos legais, sobretudo pela eliminação da necessidade de uma autoridade centralizada para verificar a autenticidade de um documento. Uma entidade emissora pode simplesmente armazenar a assinatura e a marcação de tempo associada com um documento legal na cadeia de blocos e validá-lo em qualquer tempo usando os mecanismos nativos da tecnologia [8].

Como é considerada à prova de fraudes e pode ser verificada por terceiros, de forma independente, este tipo de certificação provido por DLTs pode ser juridicamente relevante. Além disso, o registro da publicação usando *timestamp* e *hash* criptográficos de arquivos em cadeia de blocos oferece um novo e irrefutável nível de certificação. Associadamente, o uso de *blockchain* para esse tipo de registro pode ainda permitir assegurar a privacidade do documento e dos autores envolvidos, se for o caso.

Neste contexto, acreditamos que as DLTs, combinados com outros padrões emergentes de certificação e identidade digital, podem representar um caminho

interessante para permitir que documentos digitais relevantes emitidos em pu país possam ser validados em outro.

Como trabalho futuro, gostaríamos de discutir um modelo de autenticação federada para DDRs, especialmente diplomas, que permita a colaboração internacional na luta contra a fraude. Este artigo tem como objetivo divulgar a iniciativa e resultados do GT-RAP e disponibilizá-los para a comunidade.

5. Resultados Obtidos e Seu Impacto

5.1 Prototipação da Plataforma

A plataforma proposta foi totalmente implementada no período de 24 meses (de maio/2017 a abril/2019). Foram produzidas quase 20.000 linhas de código e várias tecnologias e plataformas distintas como: Bitcoin, Ethereum, OpenBadges, OpenTimeStamps, Archivematica, ChainPoint, Blockcerts, RabbitMQ, Mean Stack, Docker, Swift, etc. As seções seguintes descrevem o protótipo produzido em detalhes.

5.1.1 Arquitetura do Protótipo

O protótipo do serviço de registro, autenticação e preservação de documentos acadêmicos utiliza-se de módulos clientes e módulos servidores. Os módulos clientes são utilizados pelos usuários para acessar o serviço. Os módulos servidores fornecem os serviços demandados pelos módulos clientes através de APIs públicas. O protótipo permite o uso de diferentes formas de construção e uso de módulos clientes para acessar os serviços oferecidos.

Dessa forma, os módulos clientes usados para as operações de registro e preservação de diplomas podem ser *gateways* para o serviço proposto que operam de forma embutida e/ou integrada aos sistemas de gestão acadêmica das instituições (como o SIGAA, por exemplo) ou podem ser aplicações autônomas, construídas especificamente para esse fim. Os módulos clientes de registro e preservação são, prioritariamente, destinados para uso pelas instituições de ensino credenciadas para a utilização do serviço.

Os módulos clientes exclusivos para autenticação, por sua vez, podem ser portais interativos ou aplicações *desktop* ou *mobile*, e são destinados a usuários finais e instituições que precisam fazer, por razões variadas, a verificação da autenticidade e a prova da existência do registro de um determinado documento digital.

Os módulos clientes de registro e preservação além dos módulos clientes de autenticação tanto podem acessar a API do protótipo diretamente quanto podem fazer uso de componentes que tornam transparente essa integração, chamados aqui de agentes. São previstos dois agentes distintos: i) um que encapsula as operações de registro e preservação, e ii) outro que encapsula as operações de autenticação.

O protótipo propriamente dito é composto por um conjunto de serviços, conforme pode ser visto na Figura 3. O acesso a tais serviços é intermediado por uma *API*

RESTful, chamada *RAP Server*³⁹. A API é consumida pelos módulos clientes diretamente ou através do *Agente de Registro* e do *Agente de Autenticação*. Há ainda um *Portal de Autenticação*, o qual permite o acesso interativo para as funcionalidades de autenticação através de um *browser*, e dois componentes internos: o *DLT Broker*, que encapsula a comunicação com as diversas DLTs suportadas, e o *Curador Digital*, *middleware* que cuida da preservação digital.

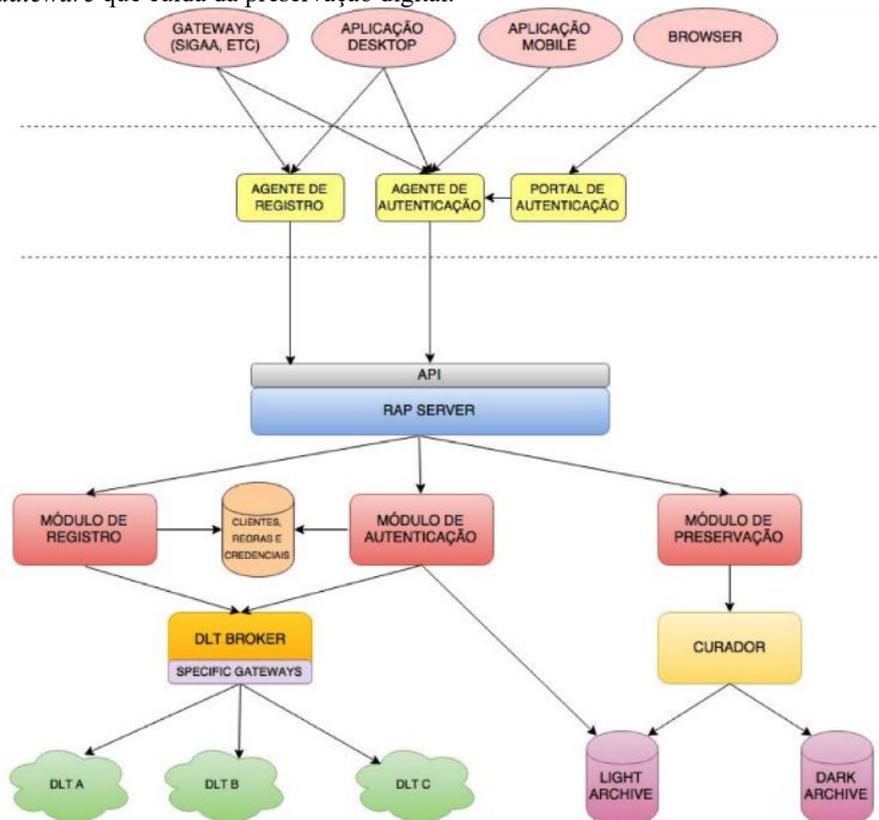
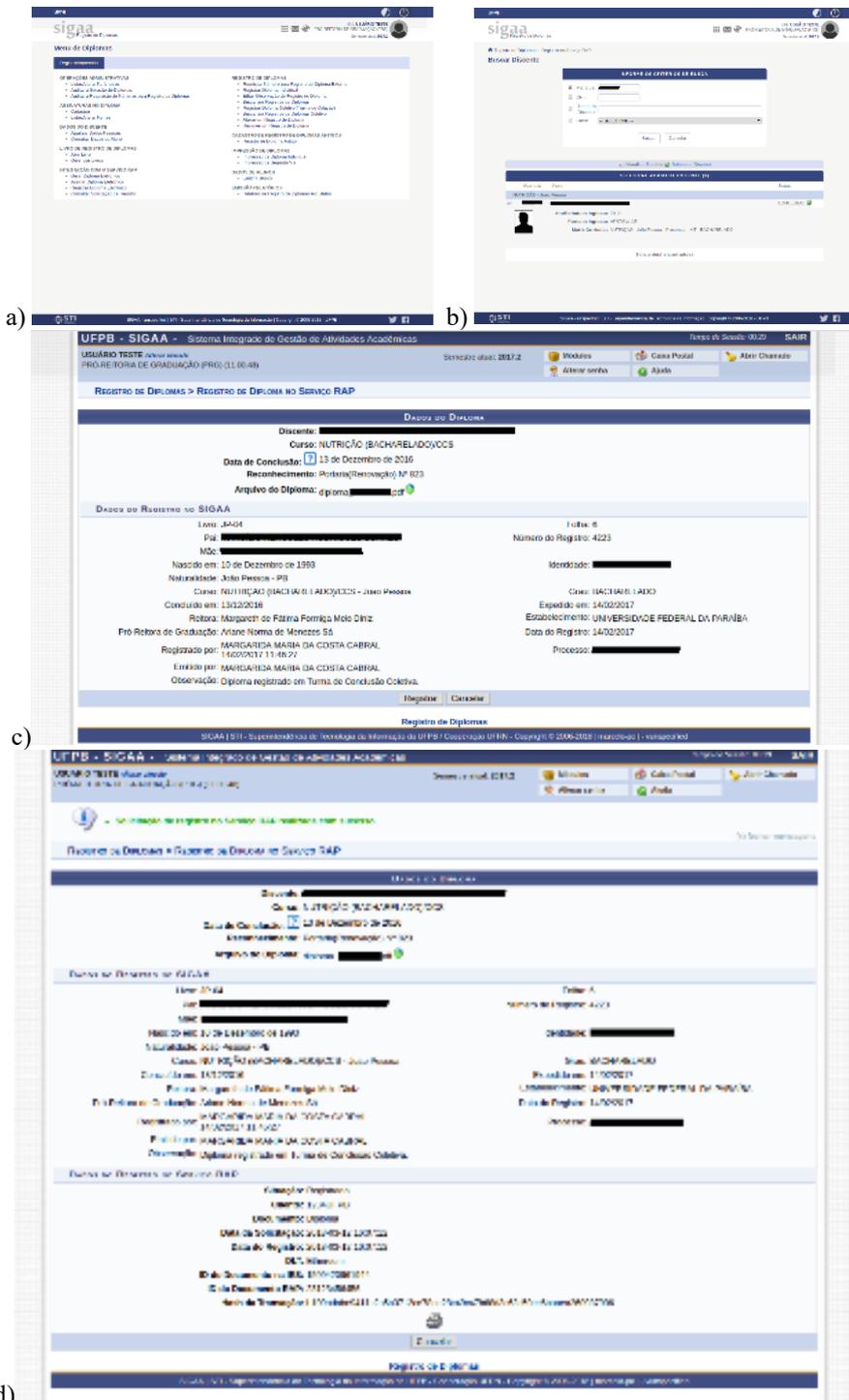


Figura 3: Arquitetura do Protótipo

³⁹ RAP é um acrônimo para Registro, Autenticação e Preservação.

9ª Conferência de Diretores de Tecnologia de Informação, TICAL2019
 “El genoma estudiantil y la metamorfosis digital universitaria”
 Cancún, México, 2 - 4 de septiembre de 2019



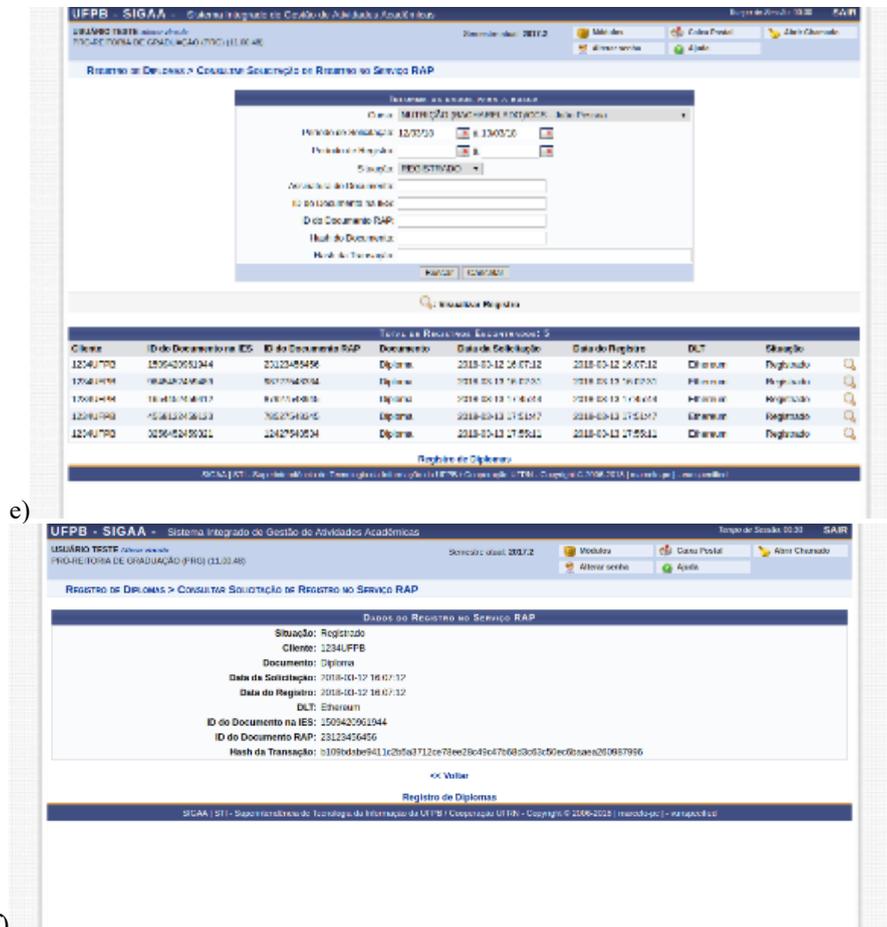


Figura 4: Telas Ilustrativas do Fluxo de Integração SIGAA/UFPB com o Protótipo RAP

5.2 Estudo Preliminar de Integração

Após o desenvolvimento do protótipo e dos agentes de registro e autenticação, o passo seguinte foi fazer uma validação preliminar da flexibilidade e adequação do protocolo de comunicação e da mensageria implementados pelo *RAP Server* e também da funcionalidade dos agentes.

As IES parceiras, UFPB e PUC-RJ, possuem ambientes e *workflows* distintos para controle da emissão de diplomas. A primeira utiliza o SIGAA, o qual já oferece um módulo para o controle de emissão de diplomas. A PUC-RJ, por sua vez, faz a gestão dos concluintes e diplomas usando um sistema específico. A presença de duas realidades bem distintas fez com que os mecanismos de integração previstos no protótipo fossem exigidos e revisados, mesmo com uma amostra pequena de IES.

No caso da UFPB, o *Módulo de Diplomas* do SIGAA ganhou um submenu

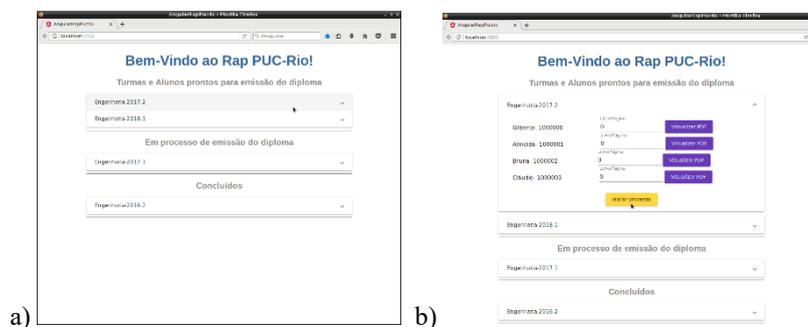
específico para integração com o protótipo como pode ser visto na Figura 4 (a). A abordagem usada na UFPB é centrada na emissão individual de diplomas, como ilustrado na Figura 2, e a solicitação de registro e preservação de diplomas é feita diploma a diploma (Figuras 4 (b), (c) e (d)). O *Agente de Registro*, o qual opera localmente, além de fazer a comunicação como o *RAP Server* para o registro de diplomas, também é ativado pelo SIGAA para realizar a geração do documento digital no formato PDF/A e também para assiná-lo digitalmente.

As solicitações de registro pendentes são acompanhadas por pedidos de status ao *RAP Server* e, quando efetivadas, são sinalizadas na base local do SIGAA, permitindo que tanto o documento digital quanto o recibo de registro na DLT sejam armazenados junto ao cadastro do egresso, assim como entregues ao interessado. As Figuras 2 (e) e (f) mostram as telas de pesquisa e recuperação da informação de registro de diplomas na DLT.

A PUC-RJ, por sua vez, possui um sistema específico para controle e emissão de diplomas e a forma de integração adotada pela equipe local foi a construção de um *gateway* para fazer a interface com o *RAP Server*. Uma outra diferença com relação à UFPB foi a abordagem baseada em turmas de formandos, tratando o registro de diplomas como operações em lote.

As Figuras 5 (a-f) ilustram o *workflow* de preparação das turmas para registro, gerando e carregando os diplomas individualmente para, só então, disparar as solicitações de registro uma a uma. Apenas quando todos os diplomas de uma turma têm os seus registros confirmados é que a turma é marcada como finalizada.

Essa é uma estratégia interessante que motivou a realização de um ajuste no protótipo para suportar a abertura e fechamento de lotes de solicitações de registro, assim como a realização de registros múltiplos na DLT através mecanismos como Árvores de Merkle para diminuir os custos, no caso de uso de DLTs tarifadas.



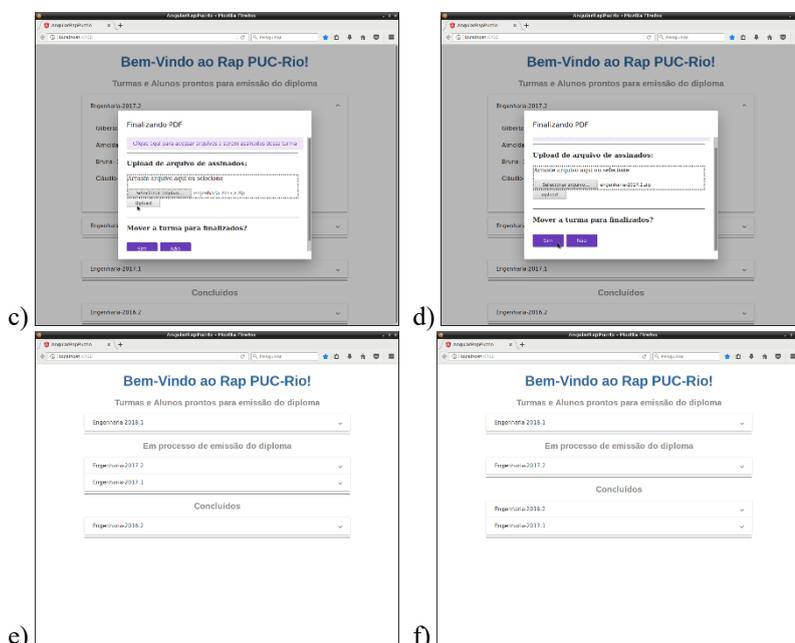


Figura 5: Telas Ilustrativas do Fluxo de Integração da PUC-RJ com o Protótipo RAP

5.3 Primeira Geração de Diplomas Digitais Baseados em Blockchain No Brasil

Os primeiros diplomas universitários digitais do Brasil, desenvolvidos com tecnologia de registro e autenticação distribuída, baseada em blockchain e certificados digitais, foram entregues no dia 21/02/2019 aos concluintes dos cursos de Ciência da Computação e Engenharia da Computação do Centro de Informática (CI) da Universidade Federal da Paraíba (UFPB) (Figura 6).



Figura 6: Cerimônia de Entrega dos Primeiros Diplomas Digitais Baseados em Blockchain na UFPB

Para simbolizar a transição histórica, os diplomas digitais foram entregues aos concluintes na cerimônia da UFPB em pen drives, embora também tenham sido enviados por email e carregados em um aplicativo especialmente desenvolvido para armazenar objetos eletrônicos dessa natureza (Figura 8b).

O template padrão de documentos digitais da Plataforma RAP usa a mesma metáfora do diploma tradicional, ou seja, uma réplica visualmente idêntica no formato padrão para preservação de longo prazo, PDF/A⁴⁰. Para permitir interoperabilidade e leitura automática dos dados do diploma, os atributos são encapsulados no arquivo PDF/A como metadados (JSON/XML) usando XMPP/Content (Figura 7).

⁴⁰ Isso deve mudar para a próxima emissão da UFPB pois o formato final adotado pelo MEC, o qual só foi divulgado após a emissão inicial, prevê um diploma no formato XML.

9ª Conferência de Directores de Tecnologia de Informação, TICAL2019
 “El genoma estudiantil y la metamorfosis digital universitaria”
 Cancún, México, 2 - 4 de septiembre de 2019

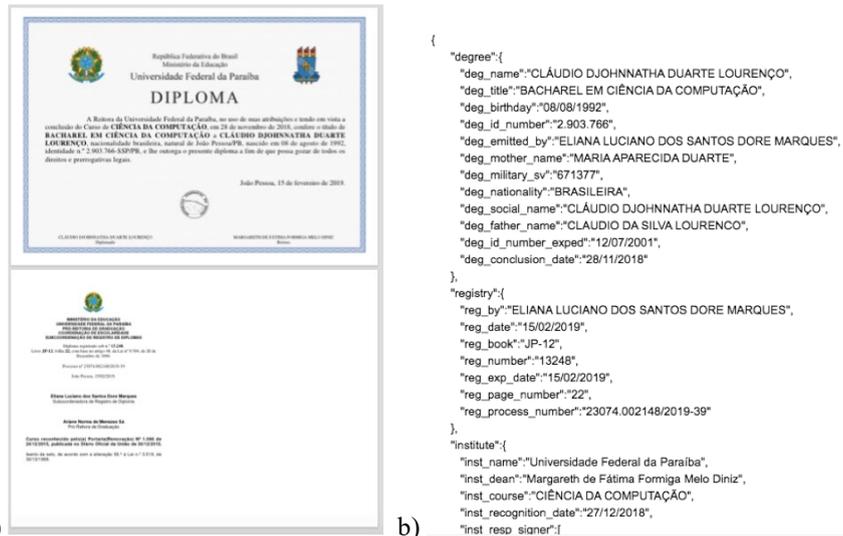


Figura 7: Representação Visual do Diploma como Arquivo PDF/A (a) com Metadados Embutidos (b) para Tratamento Automatizado

A Figura 8 mostra duas formas de validação possíveis: uma através do Portal de Autenticação da própria UFPB, o qual consome um *web service* da Plataforma RAP e através de um aplicativo mobile específico, o qual pode ser usado tanto pelo detentor do diploma quanto por um colaborador de uma empresa receptora do diploma.



a)

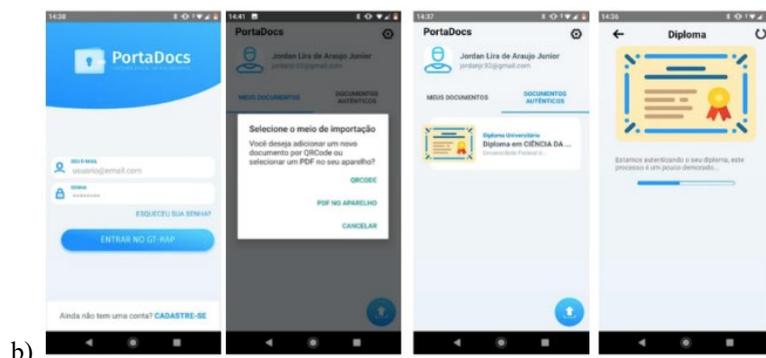


Figura 8: Validação do Diploma Digital Via Portal (a) e Aplicativo Mobile (b)

6. Lições Aprendidas

A principal contribuição científica deste projeto é a investigação de uma nova abordagem baseada no uso de DLTs e repositórios ativos distribuídos para a autenticação e preservação de longo prazo de documentos digitais legais.

Do ponto de vista tecnológico, um dos principais resultados deste trabalho foi a criação de um ambiente para dar suporte, inicialmente, a um protótipo de Serviço para Autenticação e Preservação de Documentos Digitais Acadêmicos, com potencial para adoção por diversas organizações educacionais, sejam públicas ou privadas.

Este estudo de caso específico de autenticação digital apresenta um excelente potencial de aplicação real não financeira de DLTs e pode ser o embrião para a oferta futura de um serviço permanente de grande utilidade para as IES nacionais, parcela significativa da comunidade de usuários da RNP.

O serviço proposto oferece também mecanismos que possibilitem sua integração futura com diferentes plataformas que fazem uso deste tipo de documento, como a Plataforma Lattes do CNPQ, utilizada para o cadastro de pesquisadores no país, o sistema e-MEC, utilizado pelo INEP nos processos de avaliação do ensino superior no Brasil, a plataforma Sucupira, utilizada pela CAPES no processos de avaliação da pós-graduação no Brasil e o SIGAA, sistema de controle acadêmico desenvolvido pela UFRN e amplamente utilizado por Instituições Federais de Ensino Superior.

A integração do serviço proposto com as plataformas mencionadas pode melhorar o processo de cadastramento de informações profissionais nestes sistemas, uma vez que permitirá a validação automática da autenticidade dos diplomas e certificados fornecidos, adicionando eficiência e economia ao processo e ajudando a impedir que situações como as apresentadas anteriormente voltem a acontecer.

Porém, cabe ressaltar que a infraestrutura a ser desenvolvida para dar suporte ao serviço proposto poderá apoiar também a criação de vários outros serviços para preservação e autenticação de documentos digitais e não apenas de diplomas.

Do ponto de vista da preservação digital e por conta das características de encapsulamento da replicação e do controle de falhas que pretendemos inserir na camada de armazenamento, tal abordagem pode trazer uma série de resultados complementares:

- Avançar na direção de uma integração transparente da funcionalidade de repositórios ativos com as outras camadas do modelo OAIS [12];
- Habilitar a oferta de níveis distintos de capacidade de preservação, de acordo com a relevância do objeto digital;
- Possibilidade de operar em diferentes contextos de disponibilidade e capacidade de recursos computacionais.

7. Agradecimentos

Os autores gostariam de agradecer a RNP pelo financiamento desta pesquisa através do programa de Grupos de Trabalho (GTs) e também as equipes da Superintendência de Tecnologia da Informação (STI) da UFPB e da Divisão de Admissão e Registro (DAR) e do Laboratório Telemídia da PUC-RJ pela inestimável ajuda no fornecimento de requisitos e na adaptação dos respectivos sistemas para integração com o protótipo RAP.

Referências

- [1] R. Costa, G. Lemos, V. Becker, and A. Malaguti, “Estratégias para criação de uma rede nacional para preservação digital de acervos audiovisual brasileiros,” *Reflexões sobre Preservação Audiovisual/10 anos da CineOP* â€œMostra Cinema Ouro Preto., 2015.
- [2] R. Costa, G. Lemos, V. Becker, and A. Malaguti, “We need to talk about digital preservation of audiovisual collections: Strategies for building national networks,” *jAUTI 2016: V Iberoamerican Conference on Applications and Usability of Interactive TV / 18 Convención Científica de Ingeniería Y Arquitectura.*, 2016.
- [3] FOLHA, “INEP cancela avaliação de quatro cursos por suspeita de fraude.” <http://www1.folha.uol.com.br/saber/941107-inep-cancela-avaliacao-de-quatro-cursos-por-suspeita-de-fraude.shtml>, 2011. [Online; accessed 21-March-2018].
- [4] G. do Povo, “Mp investiga o uso de mais de 500 diplomas falsos em maringá;” 2013.
- [5] G. Online, “Sedu pode levar até 180 dias para investigar diplomas falsos - mais de 100 professores processados por usarem diplomas falsos no es,” 2017.
- [6] GLOBO, “Esquema de revalidação de diploma de medicina desarticulado pela pf,” 2013.
- [7] GLOBO, “Dono de universidade denuncia esquema de venda de diplomas falsos por r\$ 550,00 em mt,” 2017.
- [8] Wikipedia, “Proof of existence.” https://en.wikipedia.org/wiki/Proof_of_Existence, 2018. [Online; accessed 21-March-2018].
- [9] M. Crosby, P. Pattanayak, S. Verma, and V. Kalyanaraman, “Blockchain technology: Beyond bitcoin,” *Applied Innovation*, vol. 2, pp. 6–10, 2016.
- [10] M. Ferreira, R. Saraiva, and E. Rodrigues, “Estado da arte em preservação digital,” 2012.

- [11] H. Kakavand, N. Kost De Sevres, and B. Chilton, "The blockchain revolution: An analysis of regulation and technology related to distributed ledger technologies," 2017.
- [12] B. Lavoie, "Meeting the challenges of digital preservation: The oasis reference model," *OCIC Newsletter*, vol. 243, pp. 26–30, 2000.
- [13] R. Ruusalepp and M. Dobrevá, "Digital preservation services: state of the art analysis. technical report, dc-net," 2012.
- [14] L. F. Sayão and L. F. Sales, "Curadoria digital: um novo patamar para preservação de dados digitais de pesquisa," *Informação & Sociedade*, vol. 22, no. 3, 2012.
- [15] K. Skinner and M. Halbert, "The metaarchive cooperative: a collaborative approach to distributed digital preservation," *Library Trends*, vol. 57, no. 3, pp. 371–392, 2009.
- [16] K. Skinner and M. Schultz, *A guide to distributed digital preservation*. Lulu.com, 2010.

Capítulo 9

Smartcampus y Tendencias - Analítica de Datos/IoT

Sistema de Registro de Asistencia basado en IoT y Tarjeta Universitaria Inteligente para la Universidad de Valparaíso, un ejemplo de Smartcampus

Marco Aravena⁴¹, Alfonso Vargas²,

¹ Universidad de Valparaíso, Prat 856, Piso 10, Valparaíso, Chile
marco.aravena@uv.cl

² Universidad de Valparaíso, Prat 856, Piso 10, Valparaíso, Chile
alfonso.vargas@alumnos.uv.cl

Resumen. La Universidad de Valparaíso es una institución chilena, estatal, pública y de una larga tradición encargada de difundir el conocimiento en distintas ramas de aprendizaje, tales como ciencias, tecnología, salud, humanidades, artes, entre otras, a través de sus 11 facultades para sus más de 18.000 estudiantes de pre y postgrado. La importancia del registro de asistencia de las clases, tanto para el académico como la universidad, se ve reflejada en diversos estudios que vinculan la inasistencia a clases con la deserción estudiantil. Los académicos cuentan con una herramienta web llamada Portal Académico, que tiene una funcionalidad que les permite crear una bitácora de clases y registrar la asistencia de los alumnos en clases. Dado que esta es una tarea que requiere necesariamente un computador y conexión a Internet, parte importante de los académicos continúan haciendo el registro de asistencia a través de una plantilla en formato papel que cual circula por los asistentes de la clase, quienes dejan en ella sus principales datos (nombre, rut y firma). Posteriormente el académico registra de los datos de la hoja de papel en el Portal Académico, siendo un proceso poco oportuno, sujeto a fallos humanos y por sobre todo inseguro, ya que el registro en papel no tiene ningún grado de autenticación por parte de los asistentes a clases.

Este trabajo presenta la implementación de una solución tecnológica de bajo costo basada en la Tarjeta Inteligente Universitaria (TUI) como herramienta de autenticación para el académico/estudiante, y dispositivos IoT (RaspBerry PI) como lectores de estas tarjetas TUI para el registro y envío de la asistencia a una base de datos central. De esta manera se facilita el registro de asistencia para el académico que ahora se realiza en forma automática e integrada a los sistemas institucionales académicos y se incorpora seguridad a través de la autenticación utilizando la TUI. El sistema desarrollado incorpora además una aplicación web para la gestión y monitoreo de los dispositivos IoT. Las pruebas del sistema a nivel de prototipo demostraron su facilidad de uso y correctitud y autenticación en el registro de los datos, contribuyendo a un registro oportuno de la asistencia para su análisis en sistema de alerta temprana de deserción.

La solución presentada abre el camino para una serie de soluciones tecnológicas basada en IoT de bajo costo y la Tarjeta TUI con el objetivo de avanzar a un campus universitario inteligente.

Palabras Clave: IoT, Asistencia, Integración, Tarjeta Universitaria Inteligente.

Eje temático: Smartcampus y Tecnologías Emergentes en la Universidad, IoT.

⁴¹ Por favor notar que todos los autores figuran con Nombre y Apellidos Completos

1. Introducción

La Universidad de Valparaíso es una institución chilena, estatal, pública, autónoma y de una larga tradición, encargada de difundir el conocimiento de las ciencias, las tecnologías, las humanidades y las artes, además de entregar competencias y valores para graduados, profesionales e investigadores comprometidos con el desarrollo regional y nacional [1]. La universidad cuenta con 11 facultades y lugares donde se imparten 42 carreras de pregrado y más de 100 programas de postgrado para los más de 18.000 estudiantes de pre y postgrado que asisten a clases en esta casa de estudios.

La Universidad de Valparaíso presenta una tasa de retención de alumnos en primer año de 81,3% para el año 2017 [2] y de 82,2% en el 2018 [3]. En la Tabla 1 se puede apreciar la comparación de todas las instituciones universitarias estatales en su tasa de retención para el año 2017.

Tabla 1. Tasas Retención 1er año para 2017 Universidades Estatales (fuente CNED Chile).

Universidad	Retención 1er Año	Nº Programas considerados
U. Arturo Prat	71,8%	81
U. De Antofagasta	74,8%	47
U. De Atacama	78,5%	20
U. De Aysen	81,7%	6
U. De Chile	83,3%	56
U. De La Frontera	82,5%	44
U. De La Serena	75,1%	40
U. De Los Lagos	79,9%	60
U. De Magallanes	79,6%	47
U. De O'Higgins	90,6%	9
U. De Playa Ancha De Ciencias De La Educación	80,4%	54
U. De Santiago De Chile	83,2%	67
U. De Talca	83,4%	39
U. De Tarapacá	80,6%	46
U. De Valparaíso	81,3%	51
U. Del Bío Bío	87,3%	40
U. Metropolitana de Ciencias De La Educación	78,7%	21
U. Tecnológica Metropolitana	80,3%	29

Considerando un ingreso de aproximadamente 3.500 alumnos para el año 2018, según la tabla Nº 1, la deserción en ese año fueron 623 alumnos. Lo anterior no sólo representa un problema para los alumnos y sus familias, si no que también afecta financieramente a la universidad por los ingresos que deja de percibir. En el ejemplo y

considerando un arancel promedio de \$ 2.000.000, la Universidad deja de percibir más de 1.200 millones por efecto de la deserción. Dado lo anterior es clave mejorar la retención con automatización de procesos orientados a la captura de información confiable y relevante como la asistencia a clases de los alumnos.

La deserción universitaria es un fenómeno que se ha venido trabajando por sociólogos y entendidos en el tema a lo largo de las últimas décadas. Los enfoques de deserción son 5 y se clasifican en psicológicos, económicos, sociológicos, organizacionales y de interacciones. El enfoque organizacional propone distintos elementos causantes, entre ellos, la experiencia de los estudiantes en la sala de clase [2, 3]. Por otro lado, las estadísticas muestran que los primeros años de universidad es donde está concentrado el menor porcentaje de asistencia, y a la vez, teniendo la asistencia una fuerte correlación con las tasas de aprobación [4]. Esto, sumado a que una de las tres principales causas de deserción entre los universitarios chilenos es su rendimiento académico [5], refleja la importancia de la asistencia en la educación universitaria en relación a la continuación y no deserción de los estudios.

La asistencia a clases es importante también, y por ende su registro, dado que la modalidad de clases de los distintos ramos impartidos por la Universidad de Valparaíso es presencial (en su gran mayoría). La asistencia a clases toma un rol principal en la vida del estudiante ya que es el método predilecto de traspaso y adquisición de conocimientos, lugar donde se tiene un feedback instantáneo por parte de los presentes y un proceso de aprendizaje compartido [6]. Si el estudiante no asiste a clases está careciendo en todo sentido de la experiencia en clases, faltando a una de las actividades académicas de mayor importancia.

Para llevar un registro de la asistencia, los académicos de la universidad cuentan con una funcionalidad en el Portal Académico, sistema web donde pueden crear una bitácora de clases con registro de asistencia para los distintos ramos impartidos. En las figuras 1 a 4 se aprecia la interfaz del Portal Académico para el registro de bitácora y asistencia.

Nº	Carrera	Perí	Seme	Nombre	Secci	Seme curríc	Estad	Fecha cierre	Bitácora asistenc
1	19085 - INGENIERIA CIVIL EN INFORMATICA - CASA CENTRAL - Dec 7263/10	2011	1	INC502 - ASIGNATURA ELECTIVA DE ESPECIALIDAD I : ADMINISTRACIÓN DE SISTEMAS	1	9	EN CURS	-	Sin alumnos
2	19083 - INGENIERIA EN INFORMACION Y CONTROL DE GESTION - CASA CENTRAL - Dec Res. Ex. 569	2011	1	IICG112 - TECNOLOGÍAS Y SISTEMAS DE INFORMACIÓN	1	1	EN CURS	-	

Fig. 1. Lista de asignaturas que dicta el académico

ICI215 - HARDWARE DIGITAL		Sección: 1		Académico: MARCO ANTONIO ARAVENA		EN CURSO		
19085 - INGENIERIA CIVIL EN INFORMATICA - CASA CENTRAL - Dec		Semestre curricular: 3		Res. Ex. N°2475				
N°	Fecha clase	Bloque	Descripción	Asistencia	Marca asister	Edit	Elimi	Agrec
		0	Caracteres restantes: 250					✓
1	23/03/2018	0	Motivación - Introducción a Digital/Análogo	13 de 15 (86.67 %)	✓	✎	✖	
2	13/04/2018	0	Redes Eléctricas	13 de 15 (86.67 %)	✓	✎	✖	
3	20/04/2018	0	Clase de ejercicios / Simulador Tinkercad	12 de 15 (80.0 %)	✓	✎	✖	
4	27/04/2018	0	Algebra de Boole / Circuitos Combinacionales	9 de 15 (60.0 %)	✓	✎	✖	

Fig. 2. Lista de clases dictadas por el académico.

✓ Marcar asistencia					✖ Cerrar
23/03/2018 - Bloque: 0					
Q buscar ...					
N°	Rut	Nombres ↕	Apellidos ↕	Asistencia	
1	19618	IGNACIO	ABARCA	✓	
2	19982	LEONARDO	ARANCIBIA	✓	
3	19616	FRANCO	ARAYA	✓	

Fig. 3. Lista de alumnos de una asignatura que dicta el académico

Bitácora asistencia ✔ Presente ✘ Ausente					
#	Fecha	Descripción	Bloque	Asistentes	Asistida
1	19/03/2018	Introducción	Sin bloque	4	SI
2	21/03/2018	Tipos de ecuaciones diferenciales. Significado y verificación de soluciones de ecuaciones diferenciales.	Sin bloque	4	SI
29	15/08/2018	Feriado legal	Sin bloque	0	NO
30	20/08/2018	Ejercitación	Sin bloque	4	SI
31	22/08/2018	III Prueba Integral	Sin bloque	4	SI
32	27/08/2018	Corrección III Prueba Integral	Sin bloque	4	SI
33	29/08/2018	Ejercitación pre Prueba Especial	Sin bloque	4	SI
Asistidas: 57.6% No asistidas: 42.4%					

Fig. 4. Vista que tiene el alumnos sobre su asistencia.

El registro de asistencia utilizando el Portal Académico requiere necesariamente un computador y conexión a internet. Existen distintos casos en que los académicos toman el registro de asistencia, a continuación se detallan los principales:

- Académico pasa lista en clases y la ingresa en el acto al Portal Académico (requiere computador y conexión a Internet).
- Académico pasa lista en clases formato plantilla papel en la que los estudiantes dejan sus principales datos (nombre y rut) y luego la traspasa en forma oportuna al portal académico.
- Académico pasa lista en clases en formato plantilla papel en la que los estudiantes dejan sus principales datos pero no la traspasa en forma oportuna o no la traspasa al portal académico.
- Académico no pasa lista en clases.

Todos los casos de registro de asistencia nombrados tienen sus desventajas, por ejemplo, una hoja dando circulando por la sala es un claro distractor mientras el académico enseña, o perder minutos valiosos al detener la clase para nombrar a cada uno de los estudiantes del curso, pero ambos métodos tienen en común el mismo problema: la carencia de autenticación del estudiante. Dado que rara vez el académico identifica a cada uno de los estudiantes de sus clases, el registro de asistencia no es del todo fiable ya que puede estar propenso a tener datos erróneos provenientes de estudiantes que registran a compañeros ausentes, fallos humanos al momento de traspasar la lista al portal, etcétera. (ver Figura 5).

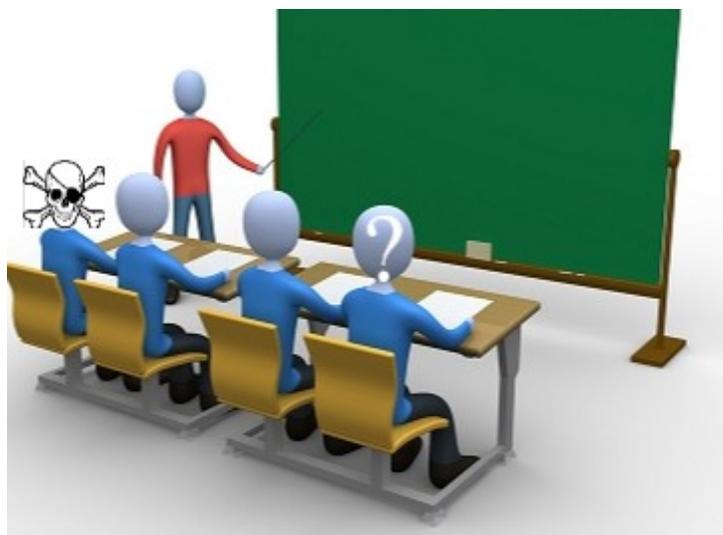


Fig. 5. Alumnos que suplantan a otros o sin autenticación

Al no tener autenticación alguna de los estudiantes, se producen datos poco fidedignos, lo que se traduce en la falta de información relevante que ayude la toma de decisiones anticipadas para evitar la deserción académica a través de un sistema de alerta temprada para la deserción.

La Universidad de Valparaíso comenzó a trabajar con la TUI (Tarjeta Universitaria Inteligente) desde el año 2011, y de ahí en adelante han ido en aumento los servicios a los cuales se pueden acceder. Esta tarjeta que es entregada a todos los estudiantes, académicos y miembros de la casa de estudios gratuitamente y tiene tecnología de identificación automática o RFID, lo que significa que actúa como una credencial electrónica segura. Dado lo anterior, se puede acceder a distintos datos de su dueño para diferenciar a cada una de las personas de la comunidad estudiantil mediante una clave de lectura sólo conocida por la institución.

Ante la necesidad de llevar un registro de asistencia con mayor grado de autenticación y de manera oportuna, se presenta en este artículo el desarrollo de una solución de bajo costo e integrada con los sistemas institucionales para la autenticación electrónica del registro de asistencia por medio de la tarjeta TUI y dispositivos IoT (en particular RaspBerry PI) como principales tecnologías. Estas herramientas fueron elegidas debido a su bajo costo y flexibilidad para proyectos con integración electrónica, obteniendo un sistema más automatizado que permite la captura y almacenamiento del registro de asistencia dentro de la sala de clases, obteniendo información que podría ser usada (entre otras utilidades) para ayudar a disminuir la deserción estudiantil en la Universidad de Valparaíso mediante la obtención de datos en un tiempo oportuno para su análisis en un sistema de alerta temprana de deserción.

2. Tarjeta Universitaria Inteligente TUI e IoT

Como se mencionó en el punto anterior, las tecnologías que componen la solución son la Tarjeta Universitaria Inteligente TUI y Raspberry Pi desde el ámbito de IoT.

La Tarjeta Universitaria Inteligente TUI es una tarjeta electrónica que se les entrega principalmente a distintas universidades nacionales e internacionales. Esta iniciativa nace en España a través del Banco Santander en el año 1995, expandiéndose y llegando a distintos países, como por ejemplo: Chile, Argentina, Brasil, Puerto Rico, México, Estados Unidos, Portugal, Polonia, Reino Unido, entre otros. Solo en Chile son más de 40 universidades e instituciones que trabajan con la TUI[7].



Fig. 6. TUI Universidad de Valparaíso.

Como se puede apreciar en la figura 6, la TUI es más que una tarjeta de plástico, contiene un chip incorporado por el medio del cual se puede acceder a distintos datos su dueño, como por ejemplo, nombre completo, rut, su rol dentro de la universidad (académico o estudiante), carrera, entre otros.

En países donde la TUI llegó a la vida universitaria hace cerca de 20 años, esta sirve como: autoservicio universitario, control de acceso, gestión de préstamo, firma electrónica, control de horario, obtención de descuentos, monedero electrónico y tarjeta de débito, entre las otras utilidades desarrolladas por las distintas universidades. Los servicios a los que se puede acceder por medio de la TUI son propios de cada universidad. En la realidad propia de la Universidad de Valparaíso se pueden encontrar diversos servicios e iniciativas, entre los que destacan:

- Autenticación para acceder a certificados oficiales. Recuperación de claves por internet.
- Préstamos de libros en las bibliotecas de la universidad.
- Reserva de salas de estudio en el Centro Integrado de Atención al Estudiante de la Universidad de Valparaíso.
- Ingreso a bibliotecas, facultad de humanidades y arquitectura, vía torniquete, dando una mayor seguridad a los estudiantes.

- AppTUI UV, aplicación con la que los estudiantes pueden acceder a información relevante además de poseer la tarjeta TUI digitalizada.

También existen una serie de planes pilotos con la tarjeta TUI, entre los que destacan: buzón de prestamos bibliotecarios, lockers inteligentes, y administración de becas de almuerzo en casinos. Además la universidad en conjunto con el banco Santander crearon el Observatorio Tecnológico TUI Chile, siendo el único en el país y en Latinoamérica, cuyo objetivo es realizar desarrollos de servicios para la TUI.

Mifare es el nombre de la tecnología con que trabaja la TUI. Esta tecnología pertenece a la empresa NXP y permite una comunicación sin contacto entre la tarjeta y un lector. Con más de 10 billones de tarjetas vendidas, los productos Mifare son hoy en día una solución en el mercado debido principalmente a su bajo costo, siendo útil en distintas aplicaciones, por ejemplo, tarjetas comerciales, de transporte, credencial electrónica, etc. La razón de éxito de esta tecnología se debe a su rapidez y al bajo costo que ella implica.

La tarjeta Mifare que se utiliza en la Universidad de Valparaiso pertenece a la familia Mifare Classic, la cual trabaja a una frecuencia de 13.56MHz con una capacidad de lectura/escritura y cumpliendo con las primeras 3 partes de la norma ISO 14443[5]. La distancia típica de lectura es en un rango menor a los 10 cm, dependiendo siempre de la potencia del lector. Las tarjetas Mifare son credenciales de memoria protegida. Estructuralmente se encuentra organizada por sectores y bloques los cuales poseen mecanismos simples de seguridad. Su capacidad de cómputo no permite realizar operaciones criptográficas de alto nivel, estando destinada principalmente a control de acceso, tarjetas de identidad corporativa, transporte urbano o ticketing.

Los sectores de memoria de la tarjeta se dividen en cuatro donde cada uno contiene de 4 a 16 bloques. Cada bloque es de 16 bytes, además el último bloque de cada sector lleva por nombre trailer y ahí se almacenan las claves de acceso llamadas 'A' y 'B', dichas claves permiten la lectura y escritura de los datos almacenados en la tarjeta.

En la figura 7 se puede observar la estructura de la tarjeta, teniendo como principales conceptos sector, bloque y trailer.

El sector 0 bloque 0 de la tarjeta corresponde al primer bloque a ser leído y contiene el identificador único de la tarjeta (UID).

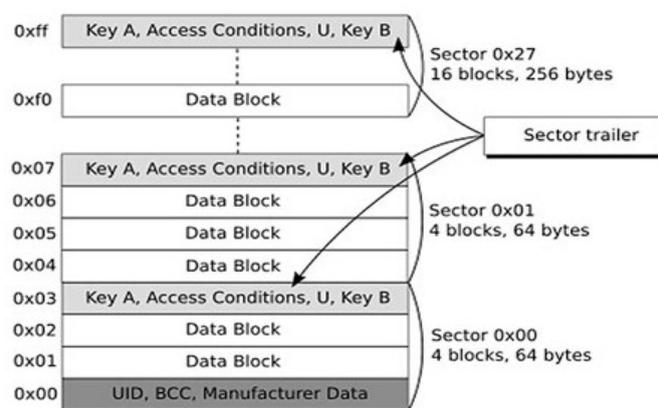


Fig. 6. Estructura Mifare.

Como se mencionó en apartados anteriores, la tarjeta necesita de un lector para poder leer los datos de identificación, lo que es propio de la tecnología RFID (Identificación por RadioFrecuencia) que corresponde a una tecnología de comunicación inalámbrica en la actúa un tag pasivo y un tag activo. Un tag pasivo no requiere una fuente de energía propio para funcionar (tarjeta TUI), en cambio el tag activo requiere su propia fuente de alimentación, lo que servirá para producir un campo electromagnético que al acercarse una tarjeta electrónica esta se inducirá eléctricamente, lo que permitirá una comunicación entre estos dos tags.

Esta tecnología nace con el misma idea que el código de barras, pero la diferencia es que en RFID no es necesario una luz infraroja para capturar los datos, además de poder almacenar una muy superior cantidad de datos que pueden ser modificados[8].

En la figura 7 se puede observar que la tecnología RFID depende principalmente de un lector que estará a la espera un tag para iniciar la comunicación.

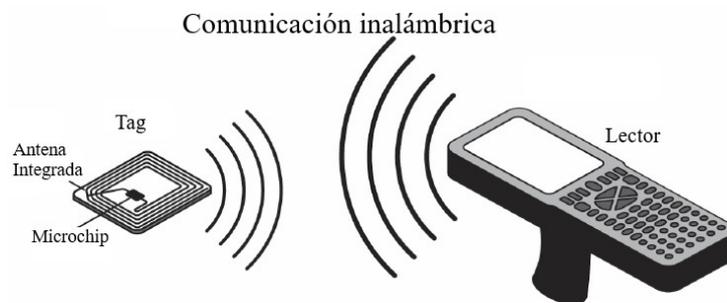


Fig. 7. Comunicación utilizando RFID.

Es común escuchar hablar de RFID y NFC (Near Field Communication) casi al mismo nivel ya que NFC deriva de la tecnología RFID. La diferencia es que la comunicación NFC es de corto alcance, lo cual limita la comunicación entre los dispositivos a un rango menor a 10 cm, lo que la hace adecuada para un intercambio más seguro de los datos. Otras diferencias es la frecuencia con que trabajan (13,56 MHz para NFC) y no es necesario emparejar los aparatos previamente (como en el caso del bluetooth) ya que la comunicación es instantánea y basta un simple toque para poder transmitir datos.[8]

Como se puede apreciar en la figura 8, la tecnología NFC destaca por su corto alcance y seguridad, además de su fácil uso ya que requiere solamente acercarse la tarjeta al lector para realizar la comunicación.

	NFC	RFID	IrDa	Bluetooth
Tiempo de Abastecimiento	<0,1 ms	<0,1 ms	~0,5 s	6 s
Alcance	10 cm	3 m	5 m	30 m
Facilidad de Uso	Fácil, intuitivo y rápido	Fácil	Fácil	Media
Seguridad	Muy alta	Buena	Visión directa	Mediante PIN
Usos	Pago, acceso, compartir información, etc...	Identificación, seguimiento, etc...	Control e Intercambio de datos	Red para intercambio de datos

Fig. 8. Comunicación utilizando RFID.

Cabe mencionar que esta tecnología no solo está limitada a las tarjetas, sino que cada vez son más las empresas que la incorporan a sus teléfonos móviles.

Para la lectura o escritura del tag pasivo, es necesario un tag activo. Este se puede encontrar en distintos formatos, tanto con entrada USB o pines para entrada de propósito general. En este desarrollo, el tag activo se implementó con un dispositivo Raspberry PI, el cual corresponde a un computador de placa reducida, lo que significa que tiene las mismas características que un computador normal, memoria, capacidad de procesamiento, conexión a internet vía cable Ethernet y WiFi, entre otras, pero que es de pequeño tamaño y bajo consumo eléctrico.

El proyecto de desarrollo de Raspberry PI nació en el Reino Unido con la intención de llevar la computación a todos los lugares posibles, debido al bajo costo del dispositivo (menor a 50 dólares), este ha tomado otro rumbo, siendo los proyectos electrónicos los que más han destacados[10]. Lo anterior se debe a que la RaspBerry PI cuenta con 40 pines que permiten conectar distintos aparatos electrónicos como LED, buzzer, lector NFC, etcétera, dando la oportunidad de crear proyectos electrónicos en distintos niveles. Y dadas sus capacidades de conectividad se convierte en un dispositivo IoT de bajo costo.

De manera general, la RaspBerry requiere de una memoria microSD para el almacenamiento, se recomienda que está sea igual o mayor a 8 GB y clase 10 para un óptimo funcionamiento. El sistema operativo recomendado para el normal uso de la RaspBerry es Raspbian, sistema operativo basado en Debian y optimizado para el hardware de la RaspBerry.

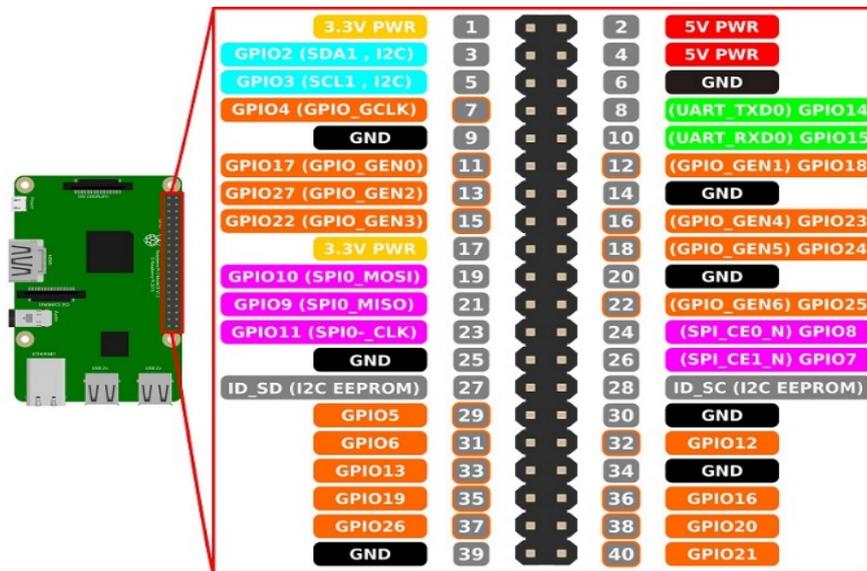


Fig. 9. Pines de interconexión de RaspBerry Pi.

Como se puede apreciar en la figura 9, la RaspBerry Pi es un computador de tamaño reducido que además de tener las mismas entradas que un computador normal (puertos USB, HDMI, Ethernet, WiFi, etc.), cuenta con GPIO Pines, que son entradas y salidas de propósito general, ideales para conectar distintos componentes electrónicos, lo que la hace una atractiva opción para la computación física y el IoT.

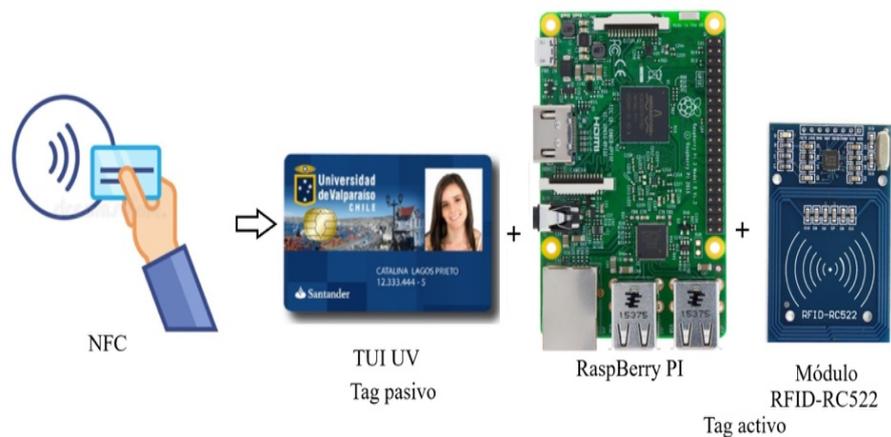


Fig. 10. Interacción Tarjeta Universitaria Inteligente TUI y RaspBerry Pi.

En la figura 10 se puede apreciar que la tecnología NFC se encuentra representada por la tarjeta TUI y la RaspBerry más un módulo de lectura conectado por medio de los pines, como lector de la TUI. La tecnología sin contacto solo requiere acercar la TUI al módulo de lectura para que los datos sean obtenidos, procesados, y almacenados por la RaspBerry.

3. Sistema propuesto para el registro de Asistencia (TUI+Raspberry PI)

Este trabajo presenta la propuesta de un sistema para el registro de asistencia basado en la tarjeta TUI como principal herramienta y método de identificación para el estudiante y académico. Esta tarjeta es leída por medio de un módulo de lectura NFC conectado a una RaspBerry PI. Este dispositivo está conectada a la red Wi-Fi de la universidad y por ese canal envía los datos obtenidos por la lector NFC hacia un servidor y bases de datos institucionales de la universidad. Se implementa también un sistema que permita la administración de los dispositivos RaspBerry por medio de una interfaz web.

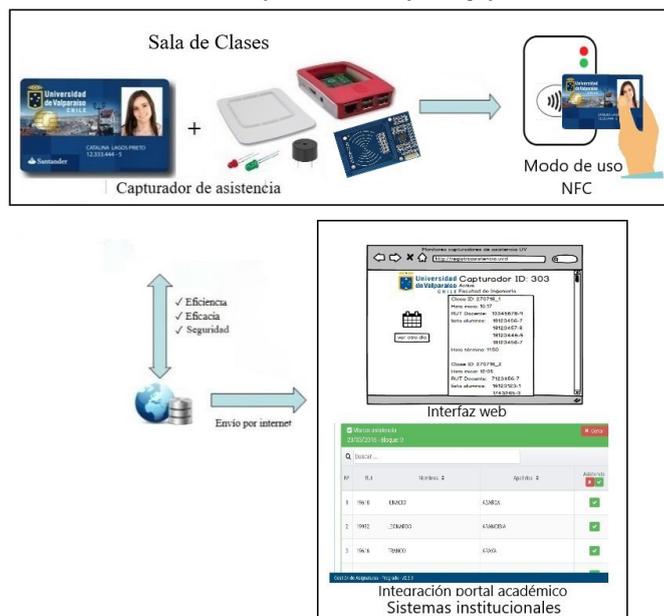


Fig. 11. Sistema de Registro de Asistencia utilizando TUI y RaspBerry Pi.

En la Fig. 11 se puede apreciar el sistema propuesto. Compuesto en una parte por la integración de Rasperry Pi + NFC con la Tarjeta TUI y por otro el sistema que recibe y administra los registro de asistencia y los dispositivos Rasperry.

Además, una característica propia de la tecnología NFC es el fácil uso, ya que solo basta con acercar la tarjeta al lector para registrarse, lo cual facilita mucho la interacción para los alumnos y académicos. De esta manera, el académico o docente inicia el

registro de asistencia acercando su TUI a la Raspberry PI para que posteriormente lo hagan todos los alumnos de la asignatura. Una vez que todos se registran, el académico cierra el registro acercando su TUI al dispositivo nuevamente. Lo anterior se puede apreciar en la Fig 12.

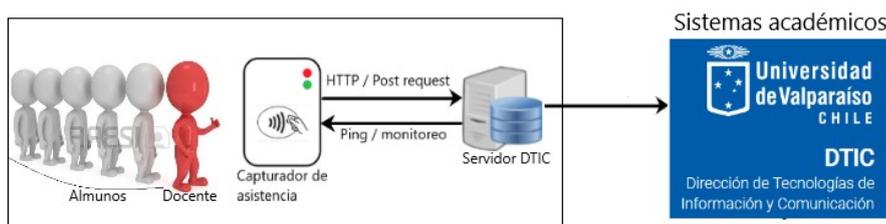


Fig. 11. Registro de asistencia en la sala de clases y envío de datos a los servidores institucionales.

Al dispositivo Raspberry Pi se le han incorporado 2 leds para indicar su estado de la siguiente manera:

- **LED Verde Permanente:** Significa que el captador de asistencia está a la espera de recibir datos provenientes de una tarjeta TUI académica, lo cual abre una clase para que posteriormente se puedan registrar los estudiantes del curso.
- **LED Verde Intermitente:** Significa que el captador de asistencia está a la espera de recibir datos provenientes de tarjetas TUI estudiantes, lo cual registra en la clase a cada uno de los estudiantes que pasaron su tarjeta después del académico. Luego que los estudiantes se hayan registrado en la clase y el académico terminado de aceptarlos en la asistencia, el académico debe volver a pasar su tarjeta TUI, lo cual termina la clase y el captador de asistencia vuelve al estado de LED verde permanente.
- **LED Rojo:** Significa que el captador de asistencia tuvo un error luego de pasar una tarjeta. Este error puede ser dado por una tarjeta no válida (una tarjeta distinta a la TUI), por pasar una tarjeta TUI no académica cuando se tiene led verde permanente, o por retirar la tarjeta TUI antes de tiempo. Después de un breve periodo de tiempo, se volverá al estado anterior del led verde.

El funcionamiento de los leds según las interacciones de los alumnos y académicos se puede ilustrar como un diagrama de estados en la Fig. 12

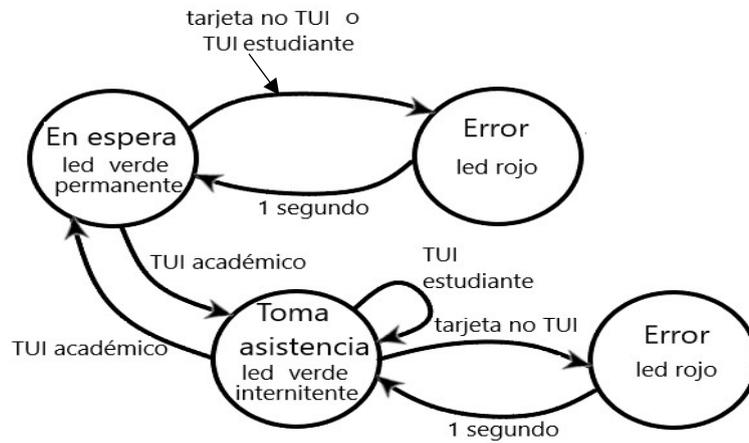


Fig. 11. Diagrama de Estados para el registro de asistencia.

Para una mejor interacción con el usuario se instaló un buzzer que emite sonidos cuando se acercan las TUI a la Raspberry Pi. La conexión física de los leds y buzzer a la Raspberry Pi se puede apreciar en la Fig. 12 y Fig 13.

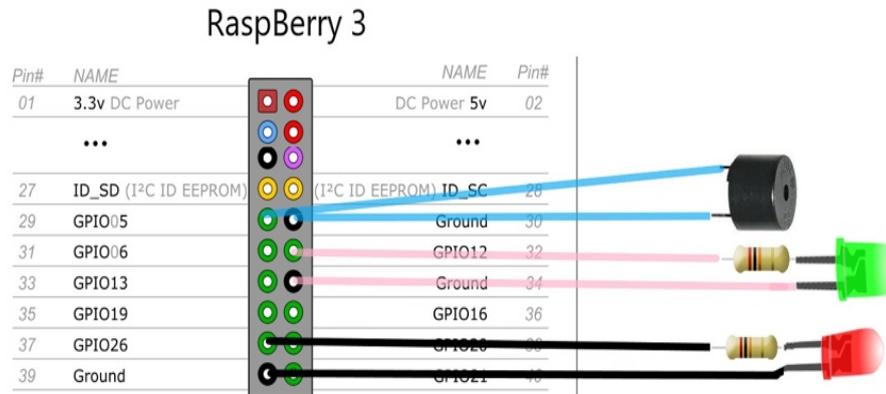


Fig. 12. Conexión física de leds y buzzer en la Raspberry Pi.



Fig. 13. Conexión física de leds y buzzer en la RaspBerry Pi (abierta y cerrada).

4. Pruebas y Resultados

Desde el punto de vista informático, se realizaron todas las pruebas requeridas, caja negra, integración etc. Sin embargo, dado que el objetivo es mejorar el proceso de captura y registro de asistencia, la prueba más relevante es la de aceptación de usuarios.

Para la prueba de aceptación de usuarios la herramienta fue puesta a prueba con el objetivo de capturar la experiencia de académicos y alumnos. Para ello realizaron pruebas alfa y beta, obteniendo buenos resultados para ambas pruebas y pudiendo rescatar información útil para la mejora del sistema.

Las pruebas fueron realizadas en un ambiente real, es decir, una sala de clases en la Facultad de Ingeniería. Dada la simplicidad de uso del capturador de asistencia, no se entregó un manual de usuario, solo una explicación a los estudiantes y académicos.

Prueba alfa

En la prueba alfa se observó cierta tendencia de los estudiantes a colocar adecuadamente la tarjeta para que fuera leída por el dispositivo NFC, por lo que se enmarcó una zona rectangular como guía del lugar de acercamiento de la tarjeta a la Raspberry PI..

En primera instancia, alrededor del 70 % de los estudiantes lograron registrarse en el primer intento y el restante 30 % logró hacerlo en segundo intento. Ningún estudiante realizó más de dos intentos para registrarse.

Prueba beta

Una vez corregido el inconveniente de la prueba alfa se realizó la prueba beta en la cual no hubo interacción entre el guía y los estudiantes.

En esta prueba se registraron 2 clases, y hubo una tendencia a poder registrarse más fácilmente ya que los estudiantes que se pudieron registrar en un primer intento fue cerca del 85 %, mientras que un solo un 15 % se registró en una segunda oportunidad. Se mantuvo la tendencia que ningún estudiante necesitara 3 o más intentos para registrarse en la asistencia.

Resultados

Se confirma la facilidad de uso del sistema dada la naturaleza de la tecnología NFC, ya que solo basta un acercamiento de la tarjeta por unos breves segundos para poder iniciar el registro de la clase o registrar la asistencia.

Las distintas fases de pruebas han servido para identificar errores y mejorar el sistema. Se realizaron cambios para facilitar aún más la interacción entre los usuarios y el capturador de asistencia. Las pruebas de mayor importancia fueron las prueba beta ya que se realizaron en un entorno real y sin ayuda del guía, comprobando que los estudiantes luego de un registro de asistencia guiada ya tienen el dominio suficiente para registrarse en la asistencia sin problema.

Con esta fase de prueba se asegura la calidad y funcionalidad del sistema, obteniendo el 100% de los datos enviados desde el capturador de asistencia hacia el servidor institucional.

5. Conclusiones

Este trabajo ha presentado el desarrollo de una solución tecnológica basada en Tarjeta Universitaria Inteligente y dispositivos IoT (Raspberry PI) en la Universidad de Valparaíso, una institución chilena, estatal, pública y de una larga tradición encargada de difundir el conocimiento en distintas ramas de aprendizaje. El desarrollo tiene por objetivo apoyar el proceso de registro de asistencia de alumnos a las clases universitarias. Lo anterior debido a la importancia del registro de asistencia de las clases, tanto para el académico como la universidad, que se ve reflejada en diversos estudios que vinculan la inasistencia a clases con la deserción estudiantil. Los académicos cuentan con una herramienta web llamada Portal Académico, que tiene una funcionalidad que les permite crear una bitácora de clases y registrar la asistencia de los alumnos en clases. Dado que esta es una tarea que requiere necesariamente un computador y conexión a Internet, parte importante de los académicos continúan haciendo el registro de asistencia a través de una plantilla en formato papel que cual circula por los asistentes de la clase, quienes dejan en ella sus principales datos (nombre, rut y firma).

La solución presentada está compuesta por el uso de la Tarjeta Inteligente Universitaria (TUI) como herramienta de autenticación para el académico/estudiante, y dispositivos IoT (RaspBerry PI) como lectores de estas tarjetas TUI para el registro y

envío de la asistencia a una base de datos central. De esta manera se facilita el registro de asistencia para el académico que ahora se realiza en forma automática e integrada a los sistemas institucionales académicos y se incorpora seguridad a través de la autenticación utilizando la TUI. El sistema desarrollado incorpora además una aplicación web para la gestión y monitoreo de los dispositivos IoT. Las pruebas del sistema demostraron su facilidad de uso y correctitud y autenticación en el registro de los datos, contribuyendo a un registro oportuno de la asistencia para su análisis en sistema de alerta temprana de deserción.

Por último, la solución presentada abre el camino para una serie de soluciones tecnológicas para un smartcampus, basada en IoT de bajo costo y la Tarjeta TUI con el objetivo de avanzar a un campus universitario inteligente.

Agradecimientos

Este trabajo ha sido financiado parcialmente por el Observatorio Tecnológico TUI Chile. <https://observatoriotui.uv.cl>

Referencias

1. Sitio Web Institucional Universidad de Valparaíso <http://www.uv.cl>.
2. Deserción universitaria en Chile: incidencia del financiamiento y otros factores asociados. <https://es.scribd.com/document/189622317/>.
3. Erika Himmel. Modelos de análisis de la deserción estudiantil en la educación superior. 2018.
4. The importance of university lecture attendance. http://www.albertosuarez.es/Doc/8Articulo_ingles.pdf
5. Estudio sobre causas de la deserción universitaria. <https://www.oei.es/historico/pdf2/causas-desercion-universitaria-chile.pdf>.
6. Showing up: The importance of class attendance for academic success in introductory science courses - the American biology teacher. [http://www.bioone.org/doi/full/10.1662/0002-7685\(2003\)065%5B0325:SUTIOC%5D2.0.CO;2](http://www.bioone.org/doi/full/10.1662/0002-7685(2003)065%5B0325:SUTIOC%5D2.0.CO;2).
7. Tarjeta universitaria inteligente. https://www.santander.com/csgs/Satellite/CFWCSancomQP01/es_ES/Corporativo/Sostenibilidad/Santander-Universidades/Tarjeta-Universitaria-Inteligente-TUI.html.
8. What is rfid?. <https://www.epc-rfid.info/rfid>.
9. What is the difference between rfid and nfc? — howstuffworks. <https://electronics.howstuffworks.com/difference-between-rfid-and-nfc.htm>
10. Eben Upton and Gareth Halfacree. Raspberry Pi user guide. 2016

Extracción de datos desde una plataforma de administración de equipos que brindan servicios de internet y su procesamiento para desarrollo de un tablero de indicadores

Rubén Yáñez Reyna^{a,b}, Cecilia Hurtado Valdez^a, Jesús Alberto Moctezuma Gómez^a, Luis Alejandro León Dávila^{a,c}

^a Área de Proyectos de Innovación, Conectividad e Inclusión Digital, Universidad de Guadalajara, Vidrio 1792, Colonia Americana, Guadalajara, Jalisco, México
ruben.yanez@redudg.udg.mx, cecilia.hurtado@redudg.udg.mx, alberto.moctezuma@redudg.udg.mx

^b Coordinación de Programas Educativos, Sistema de Universidad Virtual, Universidad de Guadalajara, Av. La Paz 2453, Colonia Arcos Sur, Guadalajara, Jalisco, México
ruben.yanez@academicos.udg.mx

^c Centro Universitario de los Valles, Universidad de Guadalajara, Carretera Guadalajara - Ameca Km. 45.5 C.P. 46600. Ameca, Jalisco, México
luis.leon@redudg.udg.mx

Resumen. El siguiente trabajo expone las experiencias en cuanto al uso de herramientas y desarrollo de métodos para extraer, procesar y mostrar información sobre el uso del servicio de internet prestado por un programa federal. La fuente de datos utilizada es la plataforma de administración de equipos de conectividad "Dashboard Cisco-Meraki", por tratarse del fabricante que provee el equipamiento a los sitios de internet que conforman la muestra. Los datos extraídos son referentes a las conexiones, así como al tráfico recibido y enviado desde los dispositivos mediante los cuales los usuarios acceden al servicio de internet. Los datos se almacenan de forma mensual, a partir de febrero de 2018 y se obtienen alrededor de 4 millones de registros de conexiones de usuario diariamente y 120 millones de registros mensuales. Se expondrá cómo se realiza la extracción, almacenamiento y procesamiento de estos datos y cómo posteriormente se interactúa con otras bases de datos que contienen variables sociodemográficas que describen los sitios de internet analizados. Lo anterior con el objetivo de realizar cruces de información para calcular algunos indicadores relacionados con los beneficios del programa federal, mismos que son presentados en un tablero de control de inteligencia de negocios.

Palabras Clave: Big data, business intelligence, extracción de datos, API, REST.

Eje temático: Smartcampus y tendencias tecnológicas emergentes en la Universidad (aplicaciones de la inteligencia artificial, blockchain, analítica, IoT, impresión 3D, experiencia inmersiva).

1. Contexto

Con el objetivo de cuantificar los beneficios de un programa federal⁴² que proporciona internet gratuito en sitios públicos, y para su visualización en un tablero de indicadores, se realizaron los convenios requeridos⁴³ para acceder a los datos generados por las conexiones de usuarios de dichos sitios de internet del programa. Estos datos son almacenados en las diversas plataformas de administración de los fabricantes de los equipos que proporcionan la conectividad en los sitios de internet del programa federal.

Inicialmente se realizó un análisis de las restricciones con las que cuentan las plataformas de los diversos fabricantes para acceder a la información; en algunas de ellas, los datos no se concentran en una nube accesible vía web, sino que se guardan en un servidor, y aunque los datos si son accesibles vía API, se requieren licenciamientos que generan costos que no desean asumir los proveedores del servicio de internet. Otras plataformas requieren que los proveedores adquieran hardware para poder realizar las consultas de los datos.

Con base en lo anterior, y tomando en cuenta que para el desarrollo de un tablero de indicadores es relevante que los datos requeridos se encuentren siempre disponibles, se decidió componer la muestra con sitios que cuentan con equipamiento Cisco-Meraki, cuya plataforma de administración permite obtener de manera ágil y estándar, información robusta para los análisis requeridos. El principal factor facilitador de esta plataforma, es que se encuentra disponible en la nube y se puede acceder vía interfaz web, con credenciales de acceso, o vía API mediante interfaces tipo REST hacia un URL en Internet, para esto solo se requiere contar con un API-Key, que es una cadena de texto que se obtiene desde la interfaz web para poder realizar solicitudes hacia la nube del fabricante Cisco-Meraki. Adicionalmente, los datos se están actualizando en tiempo real desde los dispositivos que brindan la conectividad, ubicados en los sitios de internet del programa. Finalmente, la interfaz web permite asignar etiquetas a dichos dispositivos y grupos de dispositivos, lo que agrega elementos de identificación que resultan útiles para el procesamiento.

Se considera importante la facilidad de acceso a los datos a través de interfaces estándares con fundamento en el creciente uso y aplicación del concepto de "Redes definidas por software" (SDN, por sus siglas en inglés). La coordinación que se puede dar entre los dispositivos que forman parte de una configuración centralizada permite que dichos dispositivos puedan ajustar sus parámetros para no afectar el funcionamiento de otros dispositivos cercanos y es posible proveer información de los dispositivos que conforman la red en la que se encuentran.

El único aspecto limitante en el proceso de extracción de datos desde la fuente, es decir, desde la nube de Meraki, fue el hecho de que sólo es posible realizar 5 peticiones por segundo por cada organización de dispositivos. Posteriormente explicaremos con mayor detalle cómo se subsanó esta limitante.

Con la selección de esta plataforma, la muestra de sitios con la que se cuenta para el análisis oscila entre 14,318 y 23,847 sitios de internet mensuales, mismos que generan

⁴² Los beneficios se cuantifican de forma parcial, ya que únicamente se utiliza una muestra de datos generados por usuarios del programa.

⁴³ La información utilizada es propiedad de la Secretaría de Comunicaciones y Transportes del Gobierno de México y se consulta y utiliza bajo convenio de confidencialidad.

datos sobre conexiones de usuarios, alrededor de 4 millones de registros diarios y 120 millones de registros mensuales. Los datos recabados para este trabajo integran información a partir de febrero de 2018 y hasta marzo de 2019.

Específicamente, los datos consumidos desde la plataforma de administración Cisco-Meraki son:

Tabla 1. Variables obtenidas vía API de Meraki para cada conexión de dispositivos de usuario. Elaboración propia.

Variable	Significado
Sent	Tráfico enviado por el consumo del dispositivo de usuario
Recv	Tráfico recibido por consumo del dispositivo de usuario
Id	Identificador que tiene asignado el dispositivo por su usuario
description	Descripción del dispositivo de usuario, tomado del propio dispositivo
mdnsname	Nombre de dominio del dispositivo de usuario, tomado del propio dispositivo
dhcphostname	Identificador del dispositivo de usuario que se registra en el equipo que le asigna su dirección IP para establecer conexión
Mac	Identificador del dispositivo del usuario que está ligado a su interfaz de red
Ip	Dirección IP del dispositivo del usuario
vlan	Identificador del tráfico de red del que el dispositivo de usuario participa. Útil para identificar tráfico de PAI-I o PAI-E
switchport	Identificador del puerto del switch al que está conectado el dispositivo

2. Descripción de la solución tecnológica

La obtención de los datos descritos anteriormente se realiza a través de peticiones WEB al portal especificado por Cisco-Meraki, que es hacia el cual cada dispositivo que brinda la conectividad en los sitios respectivos, envía la información. La invocación de URLs a este portal se realiza utilizando reglas específicas que consisten en el envío de parámetros que conforman el URL que se invoca. Las consultas se realizan cada hora, de 9:00 a 21:00 horas, todos los días de la semana, para obtener la mayor cantidad de datos posibles durante los diferentes horarios de servicio de los sitios de internet del programa.

La estructura con la que se guardan los datos en la nube de Meraki es sencilla, se tiene un elemento inicial que se llama Organization. Un proveedor de conectividad puede tener una o varias "organizaciones" en las que gestiona sus redes y dispositivos. La Organization contiene Networks, que son los espacios físicos en los que se instalan los dispositivos (Devices) que ofrecen la conectividad a los clientes (Clients) o para nuestros propósitos, dispositivos de usuarios del programa.

Todos los días a las 23:00 horas se realizan consultas a la nube de Meraki para verificar las organizaciones, redes y dispositivos disponibles, mismos a los que se acotará la consulta del día siguiente. Es necesario realizar esta acción puesto que nuestra muestra es dinámica y se modifica constantemente con base en la vigencia y disponibilidad de los servicios de los sitios de internet del programa. La figura 1 muestra la secuencia de peticiones en los niveles antes descritos, hasta llegar a los datos requeridos sobre las conexiones de los dispositivos de usuarios del programa.

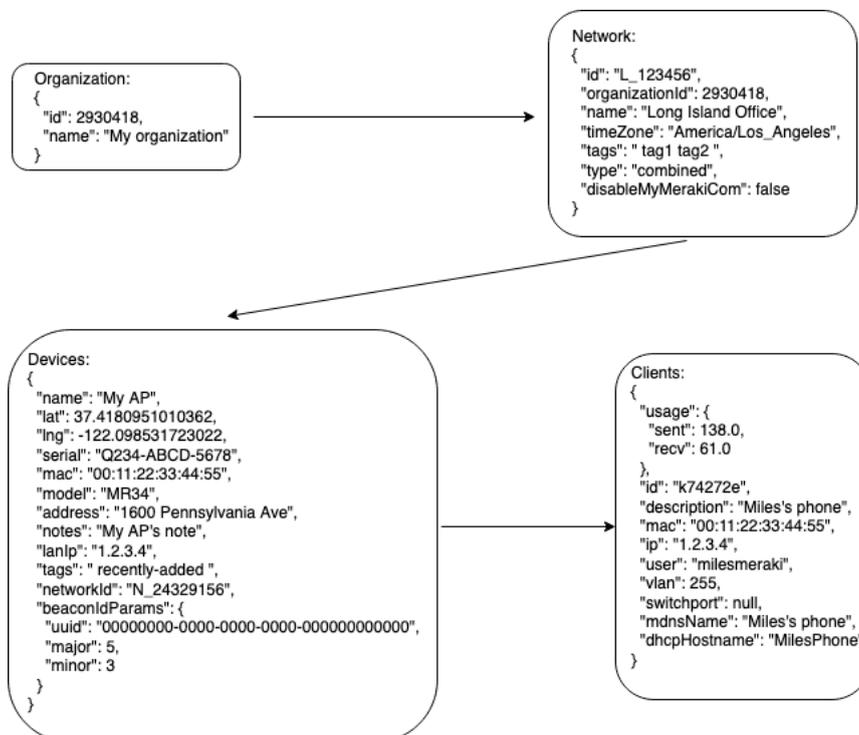


Fig. 1. Secuencia de peticiones a la nube, vía API de Meraki. Elaboración propia.

Las consultas se realizan cada hora con el fin de obtener los datos sobre las conexiones de la hora anterior, con lo que se registran datos en un periodo de tiempo diario desde las 8:00 a las 20:00 horas. Esto es así, para alcanzar a obtener los registros de conexiones y tráfico para el total dispositivos que deban ser consultados en ese día, recordando que solamente se pueden consultar 5 sitios por cada organización, por segundo. Por ello, se distribuyen el total de dispositivos a consultar a lo largo de una hora, y así sucesivamente durante 12 horas diariamente.

Se desarrollaron una serie de scripts en lenguaje python que se ejecutan de forma automática y con capacidad en multiprocesamiento, para que se realice cada hora la consulta de la información generada por los dispositivos de usuario, esto implica un proceso distinto por cada organización, que ejecuta a su vez 5 hilos o procesos adicionales cada segundo de forma paralela.

Las peticiones que se pueden realizar a la nube de Meraki para obtener los datos necesarios para la visualización de los beneficios del programa de internet gratuito, en el tablero de indicadores desarrollado, son muy simples, enseguida se muestra un ejemplo de solicitud:

```
curl -L -H 'X-Cisco-Meraki-API-Key: <key>' -X GET -H 'Content-Type: application/json' https://api.meraki.com/api/v0/devices/[serial]/clients?timespan=86400'
```

La petición "http request" es de tipo GET:

```
GET /devices/[serial]/clients
```

Los encabezados que se deben enviar en esta petición son:

```
X-Cisco-Meraki-API-Key    {{X-Cisco-Meraki-API-Key}}
```

Ejemplo de respuesta:

```
Successful HTTP Status: 200
[
  {
    "description": "Hayg's Nexus 5",
    "mdnsName": "Hayg's Nexus 5",
    "dhcpHostname": "HaygNexus5",
    "usage": {"sent": 1337.0, "recv": 7331.0},
    "mac": "00:18:D3:AD:B3:3F",
    "ip": "1.2.3.4",
    "id": "lk12uq",
    "switchport": null
  },
  ...
]
```

El periodo de tiempo en el que se requiere obtener los datos de la consulta, considerado desde el instante en que se ejecuta de la consulta hacia atrás, se debe especificar mediante el parámetro "timespan", en segundos, considerando 86,400 segundos para un día completo y con un máximo de 2,592,000 segundos de tiempo de consulta hacia atrás.

Se observa que lo que se obtiene de dicha consulta es un listado en formato JSON con la respuesta obtenida, en donde se muestran las conexiones realizadas por los dispositivos de usuarios del sitio donde se encuentra el dispositivo que brinda la conectividad.

Esta respuesta recibida en formato JSON, se transforma de manera que pueda ser almacenada en una base de datos para consultas, procesamiento y cruces de información que se explicarán posteriormente.

Los datos obtenidos desde la plataforma de administración Cisco-Meraki, son complementados para su almacenamiento con las siguientes variables:

Tabla 2. Variables que se agregan a la base de datos. Elaboración propia.

Variable	Significado
Fecha	Día, mes, año del que se está obteniendo información del dispositivo que brinda la conectividad
hora_inicio	Hora de inicio que cubre el periodo de poleo
hora_fin	Hora de fin que cubre el periodo de poleo
GID	Identificador global del sitio
NID	Identificador interno del sitio (no requerido)

Esta respuesta recibida en formato JSON, se transforma de manera que pueda ser almacenada en una base de datos para consultas, procesamiento y cruces de información que se explicarán posteriormente en la sección de aspectos críticos y relevantes.

3. Aspectos críticos y relevantes

Para gestionar los datos obtenidos mediante los métodos descritos, resultó crítico contar con un sistema de almacenamiento de datos idóneo, que cuente con un esquema que responda a los requerimientos del proyecto, pero sobre todo, que se adapte y ajuste al tipo de datos con los que se opera. En el caso de los datos descritos en este trabajo, por sus características, podrían ser almacenados en un esquema tradicional, es decir, tablas relacionadas con sus correspondientes llaves, en una base de datos que cumpla con el estándar SQL, debido a que realmente no se están almacenando datos fuera de los tipos estándares, sino que incluso en una sola tabla sería posible guardar los datos obtenidos.

La peculiaridad de nuestro proyecto se trata de la gran cantidad de datos que se obtienen de las consultas puesto que guardamos información sobre cada usuario que se conecta a los sitios de internet del programa, en la muestra de sitios utilizada. Con esta consideración se buscó una opción con características que ofrecieran velocidad de escritura; se realizaron análisis comparativos entre las bases de datos NoSQL más utilizadas, entre ellas MongoDB y Cassandra.

Los casos de uso de la base de datos Cassandra permiten ver que se ha utilizado en diversos proyectos relacionados con datos recabados vía sensores y dispositivos IoT, mensajes y datos extraídos de redes sociales, detección de fraudes y datos en series de tiempo. Adicionalmente, encontramos que su arquitectura es escalable y tolerante a fallas, su modelo de datos flexible y sus procesos de escritura y lectura son muy eficientes y permiten que aplicaciones críticas de Big Data estén disponibles de forma

continua y se puede escalar a millones de transacciones por segundo. Otra característica de Cassandra que difiere de los esquemas relacionales, es que para el modelado de datos en Cassandra se deben considerar las solicitudes (queries) que el sistema hará a la base de datos y no existen agregaciones (joins) de datos. En este esquema eficiente de consultas complejas, debe considerarse también obtener respuestas de una única tabla; esto hace que sea común que se tengan varias tablas, con datos duplicados para dar respuestas a diferentes solicitudes.

Derivado del análisis realizado, la base de datos Cassandra es utilizada para el proyecto y en este punto ha resultado muy relevante el apoyo del área de gestión de TICs de la Universidad de Guadalajara, ya que esta instancia universitaria cuenta con infraestructura de nube privada y mediante ésta se proporcionó el equipo requerido, con una gran cantidad de procesadores, así como de memoria, mismo en el que se instaló y configuró la base de datos Cassandra en un esquema de único nodo con un equipo con 90 GB de RAM y 20 procesadores.

El esquema más común de trabajo de la base de datos Cassandra contempla la configuración de un clúster de nodos en los que se puedan distribuir los procesos de almacenamiento de datos. El propio concepto de llave primaria que usa Cassandra en su modelo de datos permite ubicar cada registro en el nodo o nodos correspondientes para su consulta. Sin embargo, para este proyecto se trabajó aprovechando los recursos disponibles en el servidor que se asignó, pero ello no descarta que en un corto o mediano plazo la base de datos se extienda a un esquema de clúster para el crecimiento que se pueda requerir. La figura 2 muestra los recursos asignados al equipo en el que está ejecutándose la base de datos Cassandra.

```
top - 16:05:42 up 215 days, 3:50, 1 user, load average: 3.96, 3.45, 2.74
Tasks: 507 total, 1 running, 377 sleeping, 0 stopped, 129 zombie
%Cpu0  :  9.2 us,  3.9 sy,  0.0 ni, 83.9 id,  1.0 wa,  0.0 hi,  2.0 si,  0.0 st
%Cpu1  :  9.0 us,  2.7 sy,  0.0 ni, 87.3 id,  0.7 wa,  0.0 hi,  0.3 si,  0.0 st
%Cpu2  :  7.9 us,  3.0 sy,  0.0 ni, 87.5 id,  0.7 wa,  0.0 hi,  1.0 si,  0.0 st
%Cpu3  :  6.6 us,  6.3 sy,  0.0 ni, 85.7 id,  1.3 wa,  0.0 hi,  0.0 si,  0.0 st
%Cpu4  :  6.6 us,  3.6 sy,  0.0 ni, 89.4 id,  0.3 wa,  0.0 hi,  0.0 si,  0.0 st
%Cpu5  : 12.8 us,  2.0 sy,  0.0 ni, 84.9 id,  0.3 wa,  0.0 hi,  0.0 si,  0.0 st
%Cpu6  : 15.3 us,  5.0 sy,  0.0 ni, 78.4 id,  1.0 wa,  0.0 hi,  0.3 si,  0.0 st
%Cpu7  : 19.5 us,  4.4 sy,  0.0 ni, 75.8 id,  0.3 wa,  0.0 hi,  0.0 si,  0.0 st
%Cpu8  :  9.3 us,  2.6 sy,  0.0 ni, 87.4 id,  0.3 wa,  0.0 hi,  0.3 si,  0.0 st
%Cpu9  : 12.7 us,  1.3 sy,  0.0 ni, 85.7 id,  0.3 wa,  0.0 hi,  0.0 si,  0.0 st
%Cpu10 :  9.7 us,  0.7 sy,  0.0 ni, 89.3 id,  0.0 wa,  0.0 hi,  0.3 si,  0.0 st
%Cpu11 : 12.4 us,  2.7 sy,  0.0 ni, 84.3 id,  0.3 wa,  0.0 hi,  0.3 si,  0.0 st
%Cpu12 :  6.3 us,  3.3 sy,  0.0 ni, 89.7 id,  0.7 wa,  0.0 hi,  0.0 si,  0.0 st
%Cpu13 :  8.9 us,  3.0 sy,  0.0 ni, 87.4 id,  0.7 wa,  0.0 hi,  0.0 si,  0.0 st
%Cpu14 :  8.9 us,  2.3 sy,  0.0 ni, 88.2 id,  0.3 wa,  0.0 hi,  0.3 si,  0.0 st
%Cpu15 : 32.9 us,  6.0 sy,  0.0 ni, 60.8 id,  0.0 wa,  0.0 hi,  0.3 si,  0.0 st
%Cpu16 : 20.3 us,  3.0 sy,  0.0 ni, 76.7 id,  0.0 wa,  0.0 hi,  0.0 si,  0.0 st
%Cpu17 :  6.7 us,  2.0 sy,  0.0 ni, 91.3 id,  0.0 wa,  0.0 hi,  0.0 si,  0.0 st
%Cpu18 : 12.5 us,  2.0 sy,  0.0 ni, 85.1 id,  0.3 wa,  0.0 hi,  0.0 si,  0.0 st
%Cpu19 :  4.7 us,  7.0 sy,  0.0 ni, 87.3 id,  0.7 wa,  0.0 hi,  0.3 si,  0.0 st
KiB Mem : 93828760 total, 2664948 free, 39510772 used, 51653040 buff/cache
KiB Swap: 4194300 total, 3813716 free, 380584 used. 48327944 avail Mem
```

Fig. 2. Recursos asignados en el equipo utilizado para la base de datos Cassandra.

Los valores obtenidos del esquema establecido de único nodo para escritura, han sido de alrededor de 30 mil registros por segundo. Lo anterior aplicado a una base de

10 organizaciones de dispositivos Meraki, con alrededor de 50,000 dispositivos de conectividad⁴⁴.

Otro factor relevante en el desarrollo del proyecto es la fase de procesamiento de los datos. Las decisiones relativas a las herramientas a utilizar para el procesamiento de los datos y la realización de cálculos para obtener indicadores, se basaron en la información que se decidió mostrar como producto final en el tablero de indicadores, es decir, total de conexiones de usuarios únicos de los sitios de internet del programa, tráfico enviado y recibido en cada sitio; información desagregada tal como dispositivos conectados por sitio, por día, entre otros, así como la generación de gráficos que permiten visualizar la información en relación a clasificaciones sociodemográficas, sitios rurales o urbanos, por mencionar un ejemplo; estos datos se muestran de manera mensual.

Al final de cada mes se ha realizado el procesamiento de los datos almacenados con el fin de obtener los indicadores mencionados. Para lograr la automatización de estos procesos se utiliza Apache Spark, herramienta que permite realizar consultas a la gran cantidad de registros que se encuentran en la base de datos Cassandra y que a su vez se requieren cruzar con los datos almacenados en otras bases de datos disponibles que proporcionan información sobre características sociodemográficas de los sitios de internet del programa⁴⁵, la mayoría de estas bases de datos es de tipo relacional. Apache Spark cuenta con características que hacen que su uso e integración sea muy sencillo ya que utiliza un API unificado y resulta muy eficiente para combinar tareas de procesamiento; funciona como una especie de unificador de varios procesos sobre los mismos datos, en memoria. Los lenguajes de programación que se pueden utilizar con Spark son diversos: Java, Scala, Python. En este proyecto se seleccionó Python debido al conocimiento previo del mismo por parte del equipo, así como por la gran cantidad de bibliotecas de procesamiento de datos disponibles para lograr el desarrollo específico de los scripts requeridos para el procesamiento.

Con la interacción de Python con Spark, se desarrollaron varios programas para llevar a cabo dos principales tareas: realizar los cálculos correspondientes al conteo y discriminación de dispositivos únicos de usuarios que se conectan diariamente a los sitios de internet del programa; y ejecutar los cálculos de forma mensual para obtener los indicadores que se presentan en tablero.

El flujo de actividades que se realiza consiste en, como primer paso, descartar las conexiones a dispositivos de conectividad que no se encuentran debidamente identificados como sitios de internet del programa. Recordemos que anteriormente mencionamos que la plataforma de administración Cisco-Meraki permite identificar los dispositivos a través de etiquetas; cuando un proveedor no etiqueta adecuadamente sus dispositivos, no es posible realizar el cruce con las bases de datos con información sociodemográfica, por lo tanto esos registros no resultan de utilidad para el cálculo de indicadores. Una vez que se identifican los registros válidos, se procesan sus datos con Spark, a través de dataframes, agrupadores y funciones de agregación, y se generan dos tablas de resultados con la estructura requerida para su posterior tratamiento en la herramienta de business intelligence utilizada para la presentación del tablero de indicadores.

⁴⁴ El número de organizaciones y dispositivos de nuestra muestra es dinámico se modifica con base en la vigencia y disponibilidad de los servicios.

⁴⁵ Las bases de datos utilizadas son propiedad de la Secretaría de Comunicaciones y Transportes del Gobierno de México y se consultan y utilizan bajo convenio de confidencialidad.

La primera tabla de resultados refiere entonces a los resultados por identificador de sitio de internet. Esta es la tabla principal utilizada para calcular la mayoría de indicadores que se presentan en el tablero. Como llaves se utilizan el año, mes, identificador, entidad o región, proveedor y sector, ya que estos son los filtros disponibles para consultas desde el tablero de indicadores. En esta tabla se calcula el número de conexiones, los datos enviados, datos recibidos, número de dispositivos de usuarios del servicio de internet únicos; y se distribuye el tráfico mediante una variable con 5 categorías que permiten categorizar si los beneficios del servicio de internet se están dando hacia usuarios del PAI interno o hacia el PAI externo. A esta tabla se anexan además datos de las bases de datos sociodemográficos. Como resultados se obtiene una estructura que se muestra en la figura 3.

Column	Type	Collation	Nullable	Default
year	bigint		not null	
month	bigint		not null	
gid	character varying		not null	
estado	character varying			
proveedor	character varying			
sector	character varying			
tam_localidad	numeric			
marginacion	character varying			
tipo_poblacion	character varying			
edominloc	bigint			
latitud	double precision			
longitud	double precision			
conexiones	bigint			
datos_enviados	bigint			
datos_recibidos	bigint			
dispositivos	bigint			
ancho_banda_bajada	double precision			
ancho_banda_subida	double precision			
tipo_red	character varying			
tipo_espacial	character varying			
beneficio_pai_externo	numeric			
beneficio_pai_interno	numeric			

Indexes:

- "grouped_results_by_gid_pkey" PRIMARY KEY, btree (year, month, gid)
- "idx_grouped_results_by_gid_year_month" btree (year, month)
- "idx_grouped_results_by_gid_year_month_estado" btree (year, month, estado)
- "idx_grouped_results_by_gid_year_month_estado_proveedor" btree (year, month, estado, proveedor)
- "idx_grouped_results_by_gid_year_month_estado_proveedor_sector" btree (year, month, estado, proveedor, sector)
- "idx_grouped_results_by_gid_year_month_estado_sector" btree (year, month, estado, sector)
- "idx_grouped_results_by_gid_year_month_proveedor" btree (year, month, proveedor)
- "idx_grouped_results_by_gid_year_month_proveedor_sector" btree (year, month, proveedor, sector)
- "idx_grouped_results_by_gid_year_month_sector" btree (year, month, sector)

Fig. 3. Estructura de la tabla principal de los sitios con su identificador y sus parámetros descriptores obtenidos de otras bases de datos. Elaboración propia.

Finalmente se ejecutan los cálculos correspondientes para obtener los beneficios del programa⁴⁶.

La segunda tabla agrupa las conexiones por año, mes, entidad o región, sector, proveedor. Contando los dispositivos de usuarios únicos del servicio de internet, el tráfico enviado y recibido, se genera la estructura que se muestra en la figura 4.

⁴⁶ Para realizar estos cálculos se programaron las fórmulas definidas para calcular los elementos de medición de beneficios del programa, que se derivan de los modelos realizados para este fin y que no son objeto de este trabajo.

Column	Type	Collation	Nullable	Default
year	bigint		not null	
month	bigint		not null	
estado	character varying		not null	
sector	character varying		not null	
proveedor	character varying		not null	
usuarios	bigint			
datos_enviados	bigint			
datos_recibidos	bigint			

Indexes:

- "users_by_month_pkey" PRIMARY KEY, btree (year, month, estado, sector, proveedor)
- "idx_users_by_month_year_month_estado" btree (year, month, estado)
- "idx_users_by_month_year_month_estado_proveedor" btree (year, month, estado, proveedor)
- "idx_users_by_month_year_month_estado_proveedor_sector" btree (year, month, estado, proveedor, sector)
- "idx_users_by_month_year_month_estado_sector" btree (year, month, estado, sector)
- "idx_users_by_month_year_month_proveedor" btree (year, month, proveedor)
- "idx_users_by_month_year_month_proveedor_sector" btree (year, month, proveedor, sector)
- "idx_users_by_month_year_month_sector" btree (year, month, sector)

Fig. 4. Estructura de la tabla que integra los datos de las posibles combinaciones de filtros aplicables para mostrar resultados de usuarios y datos enviados y recibidos en un mes. Elaboración propia.

4. Resultados obtenidos

Derivado de los procesos descritos en las secciones anteriores, se obtiene como resultado la visualización de la información procesada a través de un tablero de indicadores útiles para la toma de decisiones. En este caso se optó por utilizar la herramienta de business intelligence Pentaho, a través de su módulo de Dashboard Designer. En esta interfaz se desarrollaron las vistas necesarias para mostrar los indicadores y gráficos correspondientes. El desarrollo del tablero se apoyó con algunos elementos como Java (Backend), Javascript y JQuery, así como Leaflet para mostrar algunos datos en formato de mapas de calor. El tablero cuenta con la capacidad de configuración de vistas y despliegue de datos diferenciados por rol de consulta. Las figuras 5⁴⁷ y 6 muestran las vistas de algunos indicadores del tablero y de los mapas de calor de tráfico generados, se observan también los filtros que se pueden aplicar a las consultas.

⁴⁷ Para resguardar la confidencialidad de la información, la figura muestra resultados con valor cero, con el objetivo único de ejemplificar la vista del tablero.

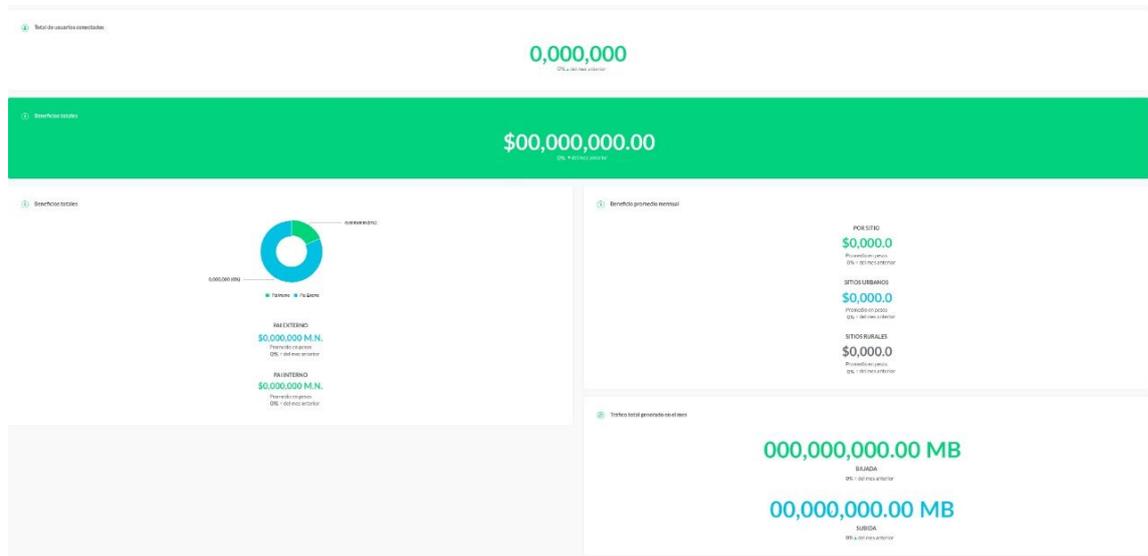


Fig. 5. Ejemplo de indicadores obtenidos en un mes a partir del procesamiento previo de los datos integrados. Elaboración propia.

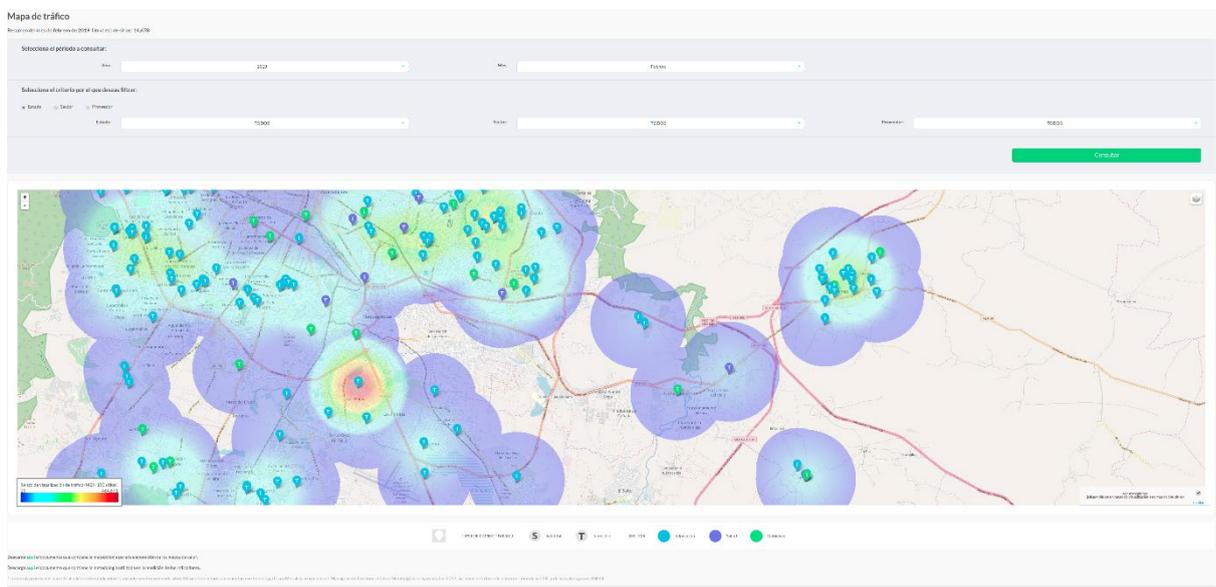


Fig. 6. Representación de la distribución de tráfico por sitio en mapas de calor, a partir del procesamiento de los datos. Elaboración propia.

El desarrollo de este proyecto resulta de gran impacto ya que establece un modelo

de negocios para recabar, analizar, procesar y presentar datos que pueden ser útiles para el seguimiento, toma de decisiones y presentación de resultados, de diversos programas públicos o privados.

5. Aprendizajes

El proceso de exploración y análisis de las diversas herramientas disponibles para realizar este tipo de proyectos, así como el establecimiento de los flujos de trabajo y modelos de negocios requeridos durante su desarrollo, deja un precedente documentado dentro de la Universidad para replicar proyectos de esta índole cuando contamos con una gran cantidad de datos y requerimos un procesamiento eficiente, expedito y con la capacidad de generar información disponible en todo momento.

Aprendemos además que cualquier herramienta utilizada tiene el potencial de ser flexible y adaptarse a las necesidades específicas de un proyecto en la medida de la creatividad de los participantes del equipo involucrados. Hacemos énfasis en que todos los componentes utilizados son de software libre, lo que elimina costos de licenciamiento y aumenta la posibilidad de integración de componentes y de expandir funcionalidades según se requiera.

Adicionalmente, este proyecto nos deja la experiencia de conformar equipos multidisciplinarios, con diversos perfiles, tales como, administradores de bases de datos, ingenieros de datos, administradores de proyecto, mismos que provienen de diversas áreas de la Universidad y que colaboran de cerca con el área de Gestión de TIC's para concretar una óptima conclusión del proyecto.

6. Impactos no esperados

El principal impacto no esperado de este proyecto está relacionado con el tiempo de desarrollo del mismo. Las mayores desviaciones de tiempo se debieron a que las primeras extracciones de datos experimentadas, se realizaron sin el conocimiento suficiente de la situación de los datos y de las capacidades de la plataforma de administración fuente, para ordenar y categorizar los datos de forma que los registros que almacenamos sean lo suficientemente válidos para que los procesos posteriores de análisis y presentación resulten eficaces.

Aún después de conocer y definir los flujos y reglas para ordenar y categorizar de mejor manera la información, a fin de realizar extracciones de datos útiles, esta tarea de gestión corrió, en este caso, por parte de los proveedores de los servicios, por lo que el tiempo que tomaron estas actividades quedó fuera del control del equipo institucional involucrado en el trabajo, provocando desviaciones en el tiempo conclusión del proyecto.

Glosario

API	Interfaz de programación de aplicaciones, se refiere a rutinas que proveen acceso a funciones de un determinado software.
API-Key	Cadena de texto generada desde la interfaz web del dashboard Meraki, que se usa para identificar a un usuario de las consultas vía API. Se liga a una cuenta de usuario y se utiliza en el proceso de autenticación.
Backend	Es la parte que se conecta con la base de datos y el servidor que utiliza dicho sitio web.
Business intelligence	Inteligencia empresarial, inteligencia de negocios o BI. Conjunto de estrategias y herramientas enfocadas al análisis de datos.
Cisco-Meraki	Empresa de TI administrada en la nube.
Dataframes	Hoja de datos o marco de datos.
Fabricante	Nombre de la empresa o marca tecnológica que crea los equipos o soluciones que brindan la conectividad, a los que se hace referencia.
GB	Gigabyte, unidad de medida de capacidad equivalente a 1000 bytes.
IoT	Internet de las cosas (en inglés, Internet of Things).
JSON	JavaScript Object Notation, (notación de objeto de JavaScript) es un formato de texto para el intercambio de datos.
Lenguaje python	Lenguaje de programación orientado a objetos.
NoSQL	Modelo de gestión de consultas en el que no solo se utiliza el lenguaje SQL.
Organización	Nomenclatura utilizada por la plataforma de administración Cisco-Meraki que se utiliza para describir el conjunto de redes que forman parte de una sola entidad organizativa, tales como una empresa o distrito escolar.
PAI externo	Servicio de conectividad que brinda el dispositivo a la comunidad.
PAI interno	Servicio de conectividad que brinda el dispositivo a la institución.
Proveedores	Empresas responsables de proveer el servicio de conectividad a Internet en los sitios del programa federal analizado.
RAM	Random Access Memory, elemento de hardware para capacidad de un sistema.
REST	La transferencia de estado representacional (en inglés representational state transfer) o REST, describe cualquier interfaz que utilice el protocolo HTTP para obtener datos o ejecutar procesos sobre datos, en formatos como XML, JSON, etc.
Scripts	Conjunto de comandos de algún lenguaje de programación

SQL	que realizan una tarea en específico. Lenguaje específico del dominio utilizado en programación, diseñado para administrar, y recuperar información de sistemas de gestión de bases de datos relacionales.
TICs	Tecnologías de la información y la comunicación. Todos aquellos recursos, herramientas y programas que se utilizan para procesar, administrar y compartir la información mediante diversos soportes tecnológico.
URL	Uniform Resource Locator (Localizador Uniforme de Recursos).
WEB	Sistema de gestión de información más utilizado para la transmisión de datos a través de internet.

Referencias

1. Cisco Systems, Inc. The Cisco Meraki Dashboard API - Cisco Meraki. Recuperado de https://documentation.meraki.com/zGeneral_Administration/Other_Topics/The_Cisco_Meraki_Dashboard_API (2018).
2. Cisco DevNet. Meraki Dashboard API - Cisco DevNet. Recuperado de <https://developer.cisco.com/meraki/api/#/rest/getting-started> (2018).
3. Hart-Davis, G. Implementación de iPads en el aula: Planificación, instalación y administración de iPads en escuelas y universidades (Deploying iPads in the Classroom: Planning, Installing, and Managing iPads in Schools and Colleges9). Barnard Castle, Durham, United Kingdom: Apres (2017).
4. Armbrust, et al. Spark SQL: Procesamiento de datos relacionales en Spark (Spark SQL: Relational Data Processing in Spark). Recuperado de <https://amplab.cs.berkeley.edu/wp-content/uploads/2015/03/SparkSQLSigmod2015.pdf> (2015).
5. Chacko. Basheer. Madhu Kumar. Capturando la procedencia de Big Data Analytics utilizando la interfaz SQL (Capturing Provenance for Big Data Analytics done Using SQL Interface). Allahabad, India: IEEE (2015).
6. Ashish Patro, Suman Banerjee. COAP: un enfoque definido por software para Home WLAN Management de un API abierta (COAP: A Software-Defined Approach for Home WLAN Management through an Open API). Recuperado de http://pages.cs.wisc.edu/~patro/papers/coap_mobiarch2014.pdf (2014).
7. Manoj R, Patil. Feris, Thia. Pentaho para análisis de Big Data (Pentaho for Big Data Analytics). Birmingham, UK: Packt Publishing Ltd. (2013).
8. Rhea, Sean. Wang, Eric. Wong, Edmund. Atkins, Ethan. Storer, Nat. LittleTable: una base de datos de series temporales y sus usos (LittleTable: A Time- Series Database and Its Uses). Recuperado de <https://meraki.cisco.com/lib/pdf/trust/lt-paper.pdf> (2017).
9. Zaharia, et al. Apache Spark: un motor unificado para el procesamiento de big data (Apache Spark: a unified engine for big data processing). New York, NY, USA: ACM DL (2016).

La implementación de las Tecnologías de la Información y Comunicación en la Universidad Ecuatoriana aseguran el análisis de datos y la toma de decisiones en procesos académicos

María Isabel Uvidia Fassler^a, Andrés Santiago Cisneros Barahona^a, Gonzalo Nicolay Samaniego Erazo^a,
Edison Patricio Villacrés Cevallos^a, Ciro Diego Radicelli García^a

^a Universidad Nacional de Chimborazo, Av. Antonio José de Sucre Km 1,5 Vía a Guano
060150 Riobamba, Ecuador
muvidia@unach.edu.ec, ascisneros@unach.edu.ec, nsamaniego@unach.edu.ec,
pvillacres@unach.edu.ec, cradicelli@unach.edu.ec

Resumen. El Sistema de Educación Superior en Ecuador desde el año 2012 requiere de un examen de ingreso llamado Ser Bachiller que le permite al aspirante escoger carreras en un orden de prioridad y determinar su ingreso a una universidad o escuela politécnica, basada en méritos y acorde a la oferta académica presentada por las instituciones de educación superior (IES). La Universidad Nacional de Chimborazo (UNACH) mediante la implementación de Tecnologías de la Información y Comunicación (TICs), además de la aplicación del descubrimiento de conocimientos en base de datos (KDD), generó un data warehouse (DW) al recopilar la información del proceso de postulación y la asignación de becas y ayudas económicas; además, mediante Minería de Datos (DM) se identificaron patrones que proporcionan información y conocimiento de la relación entre la nota de ingreso a la universidad y los estudiantes acreedores a becas de excelencia académica. Paralelamente, mediante los reportes de inteligencia de negocios, la institución tiene la oportunidad de tomar decisiones que permitan mejorar las estrategias y fortalecer los procesos académicos; garantizando un mejor desempeño académico de los estudiantes durante su formación, resultados positivos en procesos de acreditación y sobre todo asegurando la calidad en la educación de la UNACH.

Palabras Clave: Ser Bachiller, Becas, TICs, KDD, DM, BI.

Eje temático: Respondiendo a la universidad cambiante de la era digital mediante desarrollos rápidos y flexibles.

1. Introducción

Actualmente en el Ecuador, el gobierno nacional a través de la Secretaría de Educación Superior, Ciencia, Tecnología e Innovación (SENESCYT) actualizó las políticas y mecanismos de ingreso a las universidades y escuelas politécnicas con varios objetivos, incluyendo entre ellos, elevar la oportunidad de estudios a una mayor parte de la población, elevando la pertinencia y calidad académica (Guadagni, 2016); como además, lo menciona la Ley Orgánica de Educación Superior (LOES), en su artículo 63, sección "De la Admisión y Nivelación" (Ecuador, 2018), permite que la SENESCYT mediante la Subsecretaría de Acceso a la Educación Superior (SAES), lidere el proceso de admisión o ingreso a la universidad pública ecuatoriana mediante la "postulación y aceptación de cupos". Con esta lógica, se esperó que uno de los elementos de ingreso y selección sea una prueba objetiva de carácter cuantitativo que mida los conocimientos de los bachilleres postulantes a la educación superior (Guadagni, 2016).

El proceso de ingreso a las universidades y escuelas politécnicas se lo realiza mediante una evaluación de aptitudes y conocimientos que toma el nombre del examen Ser Bachiller, que rinden a nivel nacional más de 200.000 aspirantes cada semestre, obteniendo como resultado una nota o puntaje que le permite al aspirante postular o escoger carreras en todas las universidades del país. Después de haber transcurrido siete años de la implementación de éste proceso, se puede obtener varios resultados para análisis de la situación académica de los estudiantes una vez que son parte de las universidades, y en este caso de la Universidad Nacional de Chimborazo (UNACH). Los estudiantes que se encuentran cursando sus diferentes semestres de carrera tienen varios beneficios entre ellos becas y ayudas económicas, como se detalla en el Reglamento de Becas y Ayudas Económicas para Estudiantes de la UNACH, que menciona en su artículo 2. *Ámbito.* - El presente reglamento será de aplicación obligatoria para el otorgamiento de becas y ayudas económicas, a por lo menos el 10% de los estudiantes regulares, que no cuenten con recursos económicos suficientes, con alto promedio y distinción académica, los deportistas de alto rendimiento que representen a la Institución en eventos internacionales, discapacitados, y representación en concursos académicos e investigación; siempre que todos estos estudiantes acrediten niveles de rendimiento académico (UNACH, 2017).

La educación es un componente clave en el desarrollo de las naciones (Gaviria, 2012), y además, la literatura indica que el examen de ingreso ha reducido fuertemente la deserción estudiantil en la universidad ecuatoriana, mejorado sensiblemente los índices de graduación y, como expresa el Ministerio de Educación, "se evita el desperdicio de recursos" (Guadagni, 2016); por tal motivo, la UNACH ha podido recopilar información desde el 2012 que se implementa el proceso de ingreso a las universidades, datos sobre los puntajes de ingreso de los estudiantes para analizarlos y descubrir la relación que tiene el Ser Bachiller con el rendimiento académico de los universitarios. Para este análisis se utilizó el proceso de Descubrimiento de Conocimiento en Base de datos (KDD), donde Minería de Datos (DM) es el proceso medular que integra los datos de diferentes fuentes para, posteriormente, extraer un importante conocimiento, además de reportes de Inteligencia de Negocios (BI), es decir, se puede identificar información trascendente, valiosa y útil, a partir de lo cual las instituciones van a poder tomar alguna significativa decisión (Daza Vergaray, 2016), que aporte a la calidad en la educación superior.

2. Proceso de Ingreso a las Universidades y Escuelas Politécnicas

En la LOES, artículo 63, se menciona que el ingreso a las instituciones de educación superior públicas se regula a través del Sistema de Nivelación y Admisión, para todos los y las aspirantes. El sistema se rige por los principios de méritos, igualdad de oportunidades y libertad de elección de carrera o carreras e institución. [...] Las y los aspirantes que obtengan los mejores puntajes accederán a la carrera de su elección en función de la oferta disponible en las instituciones de educación superior (Ecuador, 2018). Para lo cual, el proceso se lo realiza mediante un examen sobre 1000 puntos, que evalúa el desarrollo de las aptitudes y destrezas que los estudiantes deben alcanzar al culminar la educación intermedia y que son necesarias para el desenvolvimiento exitoso como ciudadanos y para poder acceder a estudios de educación superior, que en primera instancia tomó el nombre de examen ENES (Examen para la Educación Superior), y desde el año 2016 toma el nombre de Ser Bachiller.

Una vez que el aspirante cuenta con una nota del examen Ser Bachiller, puede realizar el proceso de postulación, que significa escoger una carrera de acuerdo a su preferencia en las universidades y escuelas politécnicas públicas del país, en función de las plazas académicas habilitadas por cada una de ellas. Para el proceso de postulación el aspirante puede escoger hasta cinco opciones de carrera en un orden de prioridad y como resultado de este proceso, el aspirante acepta un cupo y lo efectiviza mediante la matriculación en una Institución de Educación Superior (IES).

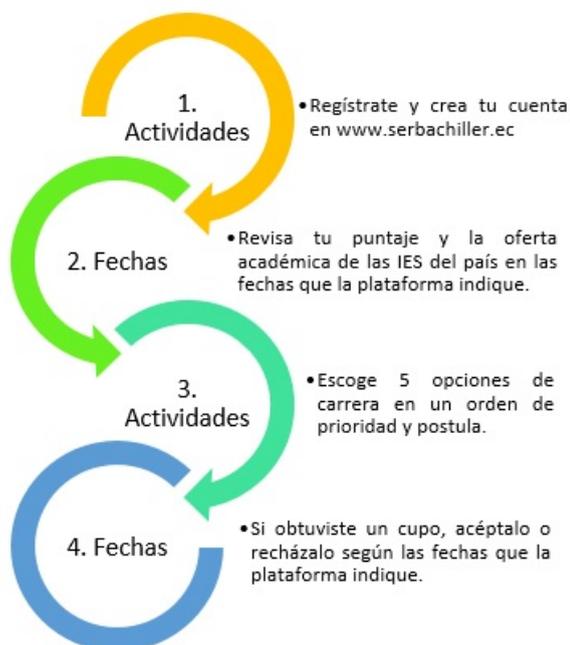


Fig. 1. Proceso de Postulación

3. Becas y Ayudas Económicas - UNACH

De acuerdo al Reglamento de Becas y Ayudas Económicas para Estudiantes de la UNACH, en su artículo Art.8. Definiciones. – menciona en los literales a. BECA. - Es la subvención total o parcial otorgada por UNACH a los estudiantes regulares que cursen sus estudios de grado en las carreras de la Universidad; y b. Estudiante Regular. - Son estudiantes regulares aquellos que se encuentran matriculados en al menos el 60% de las asignaturas, que permite su plan de estudios en el período académico ordinario correspondiente. Es decir, aquel en el que el estudiante se matriculó en todas las actividades académicas que requiere aprobar para concluir su carrera.

Además, en el artículo 10. Clases de Becas. - Se establecen las siguientes clases de becas: 1. BECAS POR RESULTADOS ACADÉMICOS: Son aquellas destinadas a los siguientes estudiantes: a. Quienes obtuvieron en el período académico inmediato anterior a la convocatoria el primero o segundo mejor promedio por paralelo de cada carrera; y, b. Quienes obtuvieron en el período académico inmediato anterior el primer, segundo o tercer puesto en concursos académicos e investigación, nacionales o internacionales en representación de la UNACH. 2. BECAS POR RESULTADOS EN EVENTOS DEPORTIVOS: Son valores económicos adjudicados a los estudiantes deportistas de alto rendimiento que representen a la Universidad y al país en eventos internacionales en las distintas disciplinas deportivas. La participación deportiva deberá haberse realizado en el período inmediato anterior a la convocatoria, contando con el Aval del Vicerrectorado Académico. 3. BECAS POR SITUACIÓN DE DISCAPACIDAD: Son valores económicos adjudicados a los estudiantes con discapacidad en condición permanente conforme las certificaciones por parte del Organismo Regulador.

4. Implementación de Tecnologías de la Información y Comunicación (TICs)

A la presente fecha, la UNACH ha implementado una herramienta informática que automatiza los procesos de selección de los becarios y que permite realizar varios reportes de análisis como resultado de este proceso importante para la comunidad universitaria. Ésta plataforma es un Sistema Académico desarrollado por la institución llamado SICOA, que consta de varios módulos a los que acceden estudiantes, docentes y personal administrativo.

Adicionalmente, se ha complementado esta solución informática con el proceso de Descubrimiento en Bases de Datos (KDD), que es un análisis automático, exploratorio y de modelado de los repositorios de datos de gran tamaño. Fayyad menciona que "El descubrimiento de conocimiento en bases de datos es el proceso no trivial de identificar patrones en datos que sean válidos, novedosos, potencialmente útiles y, por último, comprensibles" (Fayyad, Piatetsky-Shapiro, & Smyth, 1996). Complementando esta definición, los autores Maimon & Rokach, definen al KDD como un proceso organizado de identificación válida, novedosa, útil que genera patrones comprensibles de los conjuntos de datos grandes y complejos, además que este proceso KDD es iterativo e interactivo y lo define en nueve pasos (Maimon & Rokach, 2015).

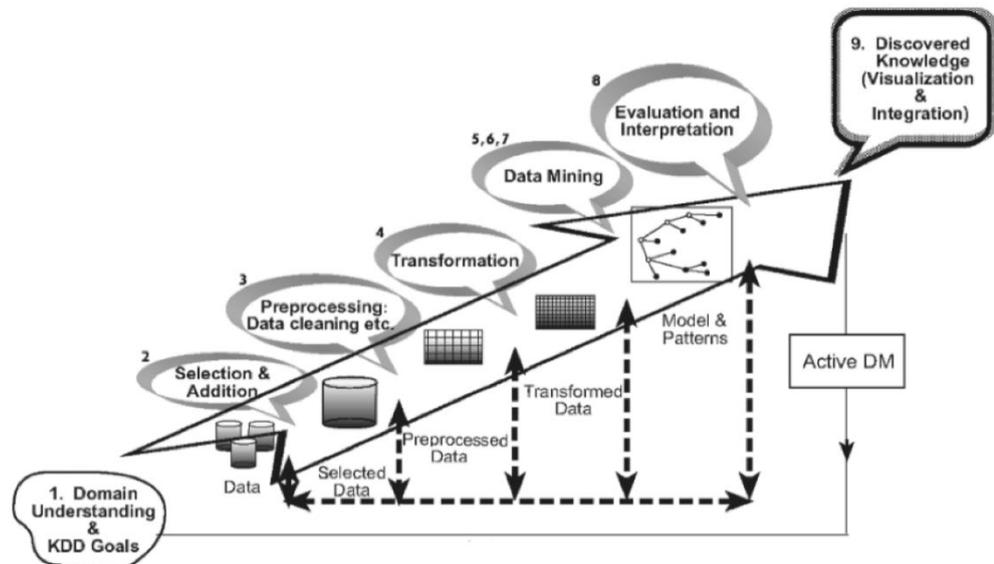


Fig. 2. Proceso de Descubrimiento de Conocimiento en Bases de Datos (KDD).

Para el proceso KDD del presente estudio, se acopló la metodología del KDD – ESPOCH aplicada a la Unidad de Nivelación y Admisión de la Escuela Superior Politécnica de Chimborazo (ESPOCH), descrita en el artículo *Moving Towards a Methodology Employing Knowledge Discovery in Databases to Assist in Decision Making Regarding Academic Placement and Student Admissions for Universities*, publicado en *Communications in Computer and Information Science*, vol 798. Springer (Uvidia Fassler, Cisneros Barahona, Ávila-Pesántez, & Rodríguez Flores, 2018). La metodología base que se contempla dentro del subproceso Data Warehouse (DW) incluye aprender el dominio de la aplicación, selección y creación de un conjunto de datos sobre la que se realiza el descubrimiento, pre-procesamiento y limpieza, y transformación de datos, con el producto DW creado con datos validados y consistentes; y en el subproceso Minería de Datos (DM) se genera la elección de la tarea de DM apropiada, elección del algoritmo DM. Empleando el algoritmo DM, se obtiene como producto los patrones generados; finalmente en los pasos: evaluación y usando el conocimiento descubierto, se consigue como producto el conocimiento que permite tomar decisiones a los usuarios académicos de la IES.

Al demostrar la utilidad y beneficio del sistema de análisis, éste fue aplicado nuevamente en la Coordinación de Admisión y Nivelación de la UNACH, con la finalidad de integrar datos que generen información y conocimiento de la relación que existe entre las notas del examen Ser Bachiller y los posibles candidatos a becarios de la UNACH. El análisis de DM fue realizado en WEKA, que es una plataforma de software libre para el aprendizaje automático y la minería de datos escrito en Java y desarrollado en la Universidad de Waikato (Waikato, 2019). Además, para complementar la visualización de los datos se utilizaron herramientas de BI, usando Qlik Sense que es una plataforma que le da superpoderes a los datos, ya que puede proporcionar los conocimientos inmediatos, explorar y crear analítica de datos

(QlikSense, 2019); para tomar las mejores estrategias que permitan asegurar la calidad en la educación superior.

5. Resultados de la Implementación de TICs

Durante la aplicación de la metodología KDD en la Coordinación de Admisión y Nivelación de la UNACH se obtuvieron los siguientes resultados:

- En el subproceso de DWH se consiguió consolidar la información de los estudiantes que obtuvieron un cupo en las carreras vigentes de la UNACH, y que fueron parte de los Cursos de Nivelación de carrera desde el año 2012 hasta el segundo periodo académico del año 2018, el cual finalizó en febrero del 2019. Además, se integró información del sistema de Becas de la UNACH, para poder generar conocimiento en base a ése análisis.
- En el subproceso de DM se obtuvieron varios resultados de análisis importantes, en los que se demuestra la relación directa entre el Ser Bachiller o nota de postulación con el rendimiento académico de los estudiantes, después de haber analizado varias técnicas, entre ellas, árboles de decisión y agrupamiento.

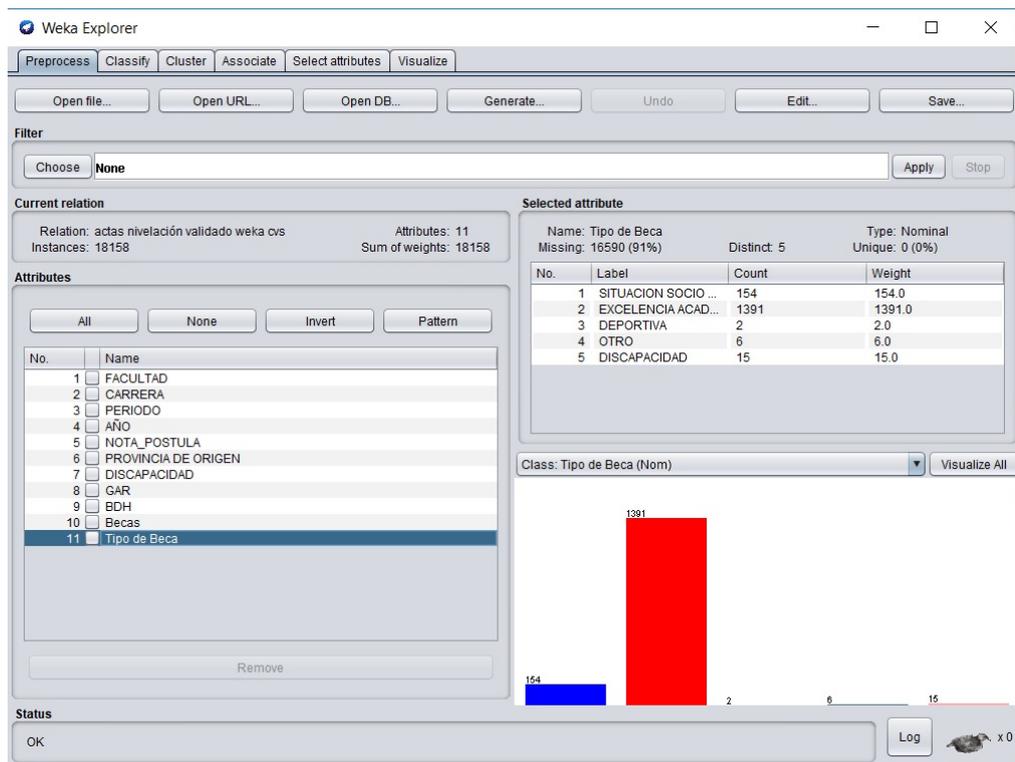


Fig. 3. DM en Weka

- Además, se realizaron reportes BI, que permiten que el proceso de toma de decisiones esté fundamentado sobre un amplio conocimiento de sí mismo y del entorno, minimizando de esta manera el riesgo y la incertidumbre. Existe una frase muy popular acerca de BI, que dice: “Inteligencia de Negocios es el proceso de convertir datos en conocimiento y el conocimiento en acción, para la toma de decisiones” (Bernabeu, 2010). Por tal motivo, a continuación, se muestran los resultados obtenidos con gran importancia para la toma de decisiones en el área académica.

Análisis de Nota Ser Bachiller - Becas

FACULTAD	AÑO	Valores	
		Avg(NOTA_POSTULA)	Count(Becas)
⊕ FACULTAD DE CIENCIAS POLITICAS Y ADMINISTRATIVAS		772,39	389
⊕ FACULTAD DE EDUCACIÓN		754,50	396
⊕ FACULTAD DE INGENIERIA		797,04	359
⊕ FACULTAD DE SALUD		858,31	424

Fig. 4. Análisis de la Nota Ser Bachiller y la cantidad de Becas asignadas por facultad

Desde la adopción del sistema de postulación y admisión a la Educación Superior Pública, la UNACH ha ofertado alrededor de 1800 cupos en cada semestre dentro de su oferta vigente que incluye 31 carreras, asociadas a 4 facultades: Ciencias de la Salud, Ciencias Políticas y Administrativas, Ciencias de la Educación e Ingeniería. Ésa es la información que se puede verificar en la tabla, en donde se analiza la nota promedio con la que los estudiantes de las facultades están accediendo a becas en la UNACH.

Fig. 5. Análisis de tipos de becas asignadas por facultad



Existen cuatro tipos de becas otorgadas por la institución, de las cuales, la beca de excelencia académica es la que se asigna con mayor frecuencia a nivel de todas las facultades, representando más del 90% de las becas totales de la UNACH.

Una vez que se conoce que la beca de Excelencia Académica es la más representativa, se enfoca el análisis nuevamente al examen Ser Bachiller, para determinar la tendencia de becarios en la UNACH, desde que ingresan y poder tomar mejores estrategias a nivel institucional sobre este importante indicador académico y de acreditación.

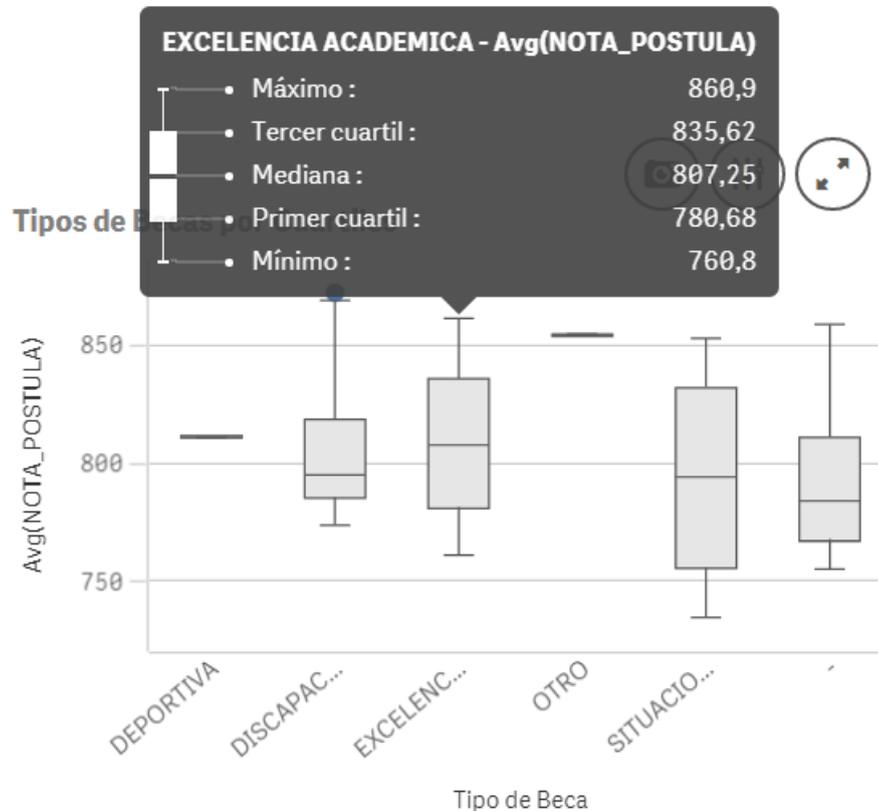


Fig. 6. Análisis de percentiles de la nota Ser Bachiller por tipo de beca

Al analizar percentiles de las notas del examen Ser Bachiller en todos los estudiantes que han aceptado un cupo en la UNACH desde el año 2012, se puede observar que, la nota máxima de beca de excelencia académica es de 860,9 puntos desde el año 2012 hasta el año 2018, la mediana representa 807,25 puntos, y que la nota mínima es de 760,8 puntos. Este análisis se lo realiza en base a los últimos siete años en los que empieza el proceso de ingreso a las universidades públicas.

Cuando se analiza el percentil de los últimos dos años de la asignación de becas, se puede identificar que en el año 2017 el valor de la mediana es de 827,59 y que en el año 2018 ésta cifra también aumenta; estos resultados están ligados totalmente a la evolución de los últimos dos años, en el que los estudiantes que están escogiendo a la UNACH como universidad para su formación de tercer nivel, tienen puntajes más altos, y al crearse mayor demanda, existe mayor competencia, y, por lo tanto, mejores puntajes.

Para complementar el análisis realizado, con el uso de gráficos de dispersión, se pudo determinar la concentración de los datos e identificar que la mayor cantidad de notas del Ser Bachiller en los estudiantes que acceden a becas de Excelencia Académica está ubicada en los rangos entre 800 y 900 puntos, permitiendo de esta forma conocer desde el inicio de la vida universitaria del estudiante, la tendencia a la hora de ser beneficiario de una beca o ayuda económica en base a la nota Ser Bachiller.

6. Conclusiones

Los retos a los cuales se está enfrentando la Universidad Nacional de Chimborazo ha brindado la posibilidad de encontrar en la implementación de Tecnologías de la Información y Comunicación, nuevas opciones, que aparte de automatizar procesos importantes para la universidad, ha permitido generar información y conocimientos con el objetivo de tener más certeza a la hora de tomar decisiones en ámbitos académicos.

Mediante la experimentación en este trabajo, se logró implementar una propuesta de la metodología de KDD para la UNACH aplicada en la Coordinación de Admisión y Nivelación, consiguiendo dentro del subproceso DW una base de datos corporativa con información histórica desde el año 2012 hasta el año 2018 de los estudiantes que postularon por las 31 carreras vigentes de la Universidad, permitiendo contar con datos para el subproceso de DM que evidencia información y conocimientos en el área académica. Dentro de esto se puede identificar la importancia de la educación media (bachillerato) y la relación en la obtención de becas de excelencia académica en la UNACH.

El conocimiento generado, permite a la UNACH contemplar al menos el 10% de la cantidad de estudiantes universitarios para el acceso a becas y ayudas económicas, que una vez que se confirma que las mejores notas de Ser Bachiller, tienden a tener becas, desde el ingreso de los estudiantes a la universidad, se debe prever la planificación de recursos económicos, que además de cumplir con el objetivo de generar igualdad de oportunidades en la educación superior, permite mejorar los indicadores de apoyo y seguimiento a los estudiantes con la finalidad de conseguir resultados positivos en procesos de acreditación, gestión de becas y aseguramiento de la calidad. Finalmente, se generan oportunidades para las carreras de las facultades con mayor cantidad de becas, ya que se asegura que el rendimiento académico de los estudiantes será aún mejor y se fortalecerá el proceso de enseñanza aprendizaje en la UNACH.

Agradecimientos

Un cordial agradecimiento a la Universidad Nacional de Chimborazo, por permitir realizar la implementación de TICS, que demuestren mejora en los procesos y sobre todo se asegure la calidad en la educación superior.

Referencias

1. Bernabeu, R. (2010). Hefesto, 146. Retrieved from <http://www.dataprix.com/data-warehousing-y-metodologia-hefesto/hefesto-metodologia-propia-para-la-construccion-un-data-wareh><http://www.dataprix.com/data-warehousing-y-metodologia-hefesto/ii-hefesto-metodologia-propia-para-la-construccion-un-data-wa>
2. Daza Vergaray, A. (2016). *Data Mining, Minería de Datos* (primera). Lima: Editorial Macro.
3. Ecuador, P. del. (2018). LEY ORGANICA DE EDUCACION SUPERIOR , LOES, 1–92.
4. Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases, 17(3), 37. <https://doi.org/10.1609/aimag.v17i3.1230>
5. Gaviria, A. (2012). Los que suben y los que bajan : educación y movilidad social en Colombia, 2012.
6. Guadagni, A. A. (2016). Ingreso a la universidad en ecuador, cuba y argentina.
7. Maimon, O., & Rokach, L. (2015). Data Mining and Knowledge Discovery Handbook. In *Data Mining and Knowledge Discovery Handbook* (pp. 22–38). https://doi.org/10.1007/0-387-25465-x_2
8. QlikSense. (2019). Qlik Sense.
9. UNACH, U. N. de C. (2017). Reglamento de becas y ayudas económicas para estudiantes de la Universidad Nacional de Chimborazo, (0098), 1–12.
10. Uvidia Fassler, M. I., Cisneros Barahona, A. S., Ávila-Pesántez, D. F., & Rodríguez Flores, I. E. (2018). Moving Towards a Methodology Employing Knowledge Discovery in Databases to Assist in Decision Making Regarding Academic Placement and Student Admissions for Universities. In M. Botto-Tobar, N. Esparza-Cruz, J. León-Acurio, N. Crespo-Torres, & M. Beltrán-Mora (Eds.), *Technology Trends* (pp. 215–229). Cham: Springer International Publishing.
11. Waikato, T. U. of. (2019). WEKA. Retrieved from <https://www.cs.waikato.ac.nz/ml/weka/>

Capítulo 10

Transformando los servicios hacia la mejora de la experiencia del usuario

Gestión del conocimiento para la atención a usuarios y soporte de servicios de TI en la Universidad Autónoma de Yucatán

Autor: Carmen Humberta de Jesús Díaz Novelo

Resumen. La Universidad Autónoma de Yucatán, UADY, cuenta con veinticinco mil estudiantes distribuidos en cinco campus y cerca de cuatro mil empleados académicos y administrativos; para brindar los servicios de Tecnologías de Información (TI), utiliza una red de voz y datos, llamada RIUADY y comprometida con la calidad de los servicios, la UADY ha certificado su proceso de provisión y soporte de TI en la norma ISO 9001:2015, sin embargo brindar los servicios de TI plantea grandes retos para la Institución, dentro de los que se encuentran: 1. crecimiento exponencial en la demanda y calidad de los Servicios de TI, 2. Saturación, alta rotación y falta de profesionalización del personal de TI; 3. tiempo limitado o nulo para la asimilación de tecnologías existentes y nuevas y; 4. el conocimiento no se registra de manera estructurada. Se identificó que existe una gran cantidad de conocimiento tácito que se desaprovecha, por ello, el objetivo del trabajo es presentar un modelo de gestión de conocimiento, que considera e integra un conjunto de elementos para sistematizar los procesos de transferencia de conocimientos y aprovecharlos en una atención a usuarios más satisfactoria y brindando el soporte de los servicios de TI más eficientemente a través de la integración del conocimiento explícito y tácito del personal, con el apoyo de las tecnologías de información. Los resultados obtenidos son: reducción de tiempos de integración del personal, reducción del tiempo de soporte a usuarios para problemas comunes, servicios estandarizados y a pesar de la demanda exponencial y crecimiento de los servicios de TI se han logrado mejorar los niveles de atención y calidad de servicio con el mismo número de personas. El impacto de este proyecto se ha visto en la mejora de la satisfacción con los servicios de TI por parte de los profesores, investigadores y alumnos que se refleja en la evaluación de los servicios. En el documento se presentan las lecciones aprendidas en la aplicación de las buenas prácticas de servicios de TI, así como las limitaciones y dificultades para la aplicación del modelo en otra Institución de Educación Superior.

Palabras Clave: Mesa de servicio, Gestión de conocimiento, ITIL, atención a usuarios, buenas prácticas.

Eje temático: Transformando los servicios hacia la mejora de la experiencia del usuario (comunidad universitaria).

1. Introducción

La Universidad Autónoma de Yucatán, UADY, es una institución de educación superior que ha avanzado en el uso de las Tecnologías de Información y la generación de conocimiento, los cuales pone a disposición de la sociedad mediante los servicios digitales que ofrece, apoyando sus funciones sustantivas de docencia, investigación y extensión. Las exigencias del entorno y el compromiso con la calidad educativa, han llevado a la UADY a contemplar la adopción de buenas prácticas, estándares y normas de calidad para la provisión de sus servicios de TI, por lo que cuenta con un proceso de provisión de tecnologías de información que se certificó en la norma de calidad ISO 9001:2008 y que en el mes de febrero del año 2019 actualizó esta certificación a la norma ISO 9001:2015.

Uno de los objetivos del sistema de gestión de la calidad es el incremento en la satisfacción de los usuarios, la UADY cuenta con veinticinco mil estudiantes distribuidos en cinco campus y cerca de cuatro mil empleados académicos y administrativos; que usan los servicios de Tecnologías de Información de la Red Integral de la UADY (RIUADY), los usuarios identificados son:

1. Rector. Máxima autoridad universitaria.
2. Directivos. Son los directores de las Facultades y Escuelas y; Directores Generales de la Institución.
3. Académicos e investigadores. Profesores e investigadores de tiempo completo, así como por contrato que laboran en la Institución.
4. Administrativos. Personal de las Facultades que se desempeña en las áreas de contabilidad, servicios escolares o como secretarías y el personal que labora en la administración Central de la Institución.
5. Alumnos. Estudiantes de las distintas Facultades en cualquier grado.
6. Sociedad en general. Personas que ingresan a los Campus para solicitar algún servicio previa autorización de algún directivo.

2. Problema

Se ha presentado un crecimiento exponencial en la demanda de servicios y los usuarios exigen una mejora en la calidad de los Servicios de TI. Al realizar el diagnóstico para la atención de la problemática se identificaron las siguientes causas:

1. Saturación, alta rotación y poca capacitación para el personal de TI.
2. Falta de Integración de las TI, proyectos aislados y duplicación de funciones del personal.
3. Tiempo limitado o nulo para la asimilación de tecnologías existentes y nuevas.
4. El conocimiento no se registra de manera estructurada, lo que ocasiona una pérdida total o parcial del mismo, existiendo ocasiones en las que es necesario empezar desde cero para generar lo que ya se tenía.
5. Existe una gran cantidad de conocimiento tácito que se desaprovecha o se pierde por no contar con estrategias de transferencia tecnológica.
6. La necesidad de disponer de estos conocimientos en algún momento y de forma inmediata

7. Los métodos tradicionales de organización de la información no permiten una administración eficiente por el exceso de información y conocimiento que se va generando.
8. Existe la necesidad de contar con nuevas metodologías, herramientas y criterios que permitan una adecuada organización de la información y que faciliten la administración del conocimiento.

3. Justificación

Se observa la necesidad de estructurar la enorme cantidad de información que se ha generado desde que la RIUADY inicia sus operaciones, en el año 1993 y la que se va generando día a día, de tal forma que pueda ser consultada en el menor tiempo posible. Se requiere registrar las experiencias y lecciones aprendidas y que esto sirva para enriquecer las nuevas experiencias.

Se puede impulsar la transformación digital de la Institución con el uso de nuevas herramientas, con las bases de conocimiento disponibles a los alumnos y profesores desde cualquier punto del mundo. En el área administrativa, un modelo de gestión de conocimiento puede permitir una forma más eficiente e intuitiva de generar y consultar procedimientos y métodos de trabajo, ahorro en papel, depuración de errores comunes, auto-aprendizaje en diversas áreas, entre muchos otros aspectos.

El modelo de gestión de conocimiento permitiría a la UADY un ahorro sustancial en recursos humanos y materiales, contribuiría en la calidad de sus programas educativos en aula y a distancia, elevaría también el desempeño de sus áreas administrativas y sentaría las bases para nuevos desarrollos a futuro con una adecuada organización del conocimiento. [1]

4. Objetivos

El objetivo general trabajo es presentar el modelo de gestión de conocimiento para la atención a usuarios y soporte de servicios de TI, desarrollado en la Universidad Autónoma de Yucatán, describiendo los elementos que lo conforman.

Objetivos particulares

1. Describir el proceso de evolución de la gestión del conocimiento en la Red integral de la UADY, RIUADY.
2. Presentar la metodología para la implementación de las Bases de conocimiento y describir su uso.
3. Tipificar y priorizar la información técnica para una asimilación más efectiva del conocimiento del personal involucrado en los procesos de Tecnologías de Información.
4. Identificar de manera efectiva la naturaleza de las diferentes problemáticas asociadas al funcionamiento y operación de los servicios de TI.

5. Describir el esquema de seguimiento granular y multinivel para la incorporación de nuevo personal.

5. Revisión de la literatura

Administración del conocimiento

El conocimiento es cualquier recurso intangible de un negocio que ayuda a su gente a hacer algo mejor de lo que podrían hacerlo sin ello. Identifica los conocimientos en una organización de la forma siguiente [2]:

Conocimiento Tácito (Capital humano) – habilidades, competencias, saber cómo (know-how) y conocimiento contextual en la cabeza de las personas (algunos autores incluyen las emociones).

Conocimiento Explícito (Capital Estructural) - Es el conocimiento que es capturado o codificado en bases de conocimiento de la compañía, herramientas, catálogos, directorios, modelos, procesos y sistemas.

Conocimiento de cliente (Capital cliente) – Es el conocimiento colectivo acerca de los clientes de la compañía, su gente, sus necesidades, hábitos de consumo, etc.

Conocimiento Innovador (Capital de innovación) – conocimiento colectivo sobre mercados no explorados o inexplorados, tecnologías, productos y procesos operativos.

En términos de un modelo general de administración de conocimiento (AC), un modelo descriptivo propuesto integra el conocimiento explícito, el conocimiento tácito y la infraestructura. El conocimiento explícito y el conocimiento tácito tienen una relación simbiótica donde el conocimiento tácito contribuye al conocimiento explícito y viceversa. Para hacer efectiva la AC es absolutamente esencial llevar juntos el conocimiento explícito y el conocimiento tácito en una infraestructura. El tipo de infraestructura depende de la complejidad de la compañía y de la disponibilidad de los recursos para soportar la AC, de las metas y objetivos de la compañía.

El capital intelectual podría ser cualquier activo que no pueda medirse pero que sea utilizado por una empresa para su beneficio. El conocimiento, la experiencia colectiva, la buena voluntad, el valor de la marca y los beneficios de patentes no aparecen directamente en documentos contables convencionales. No es de extrañar que muy pocas empresas (como Microsoft con \$ 14 miles de millones en ventas) con los niveles más altos de activos de conocimiento intelectual e intangible nunca Llegue a los escalones superiores de la lista de clasificación basada en las ventas de *Fortune*. Las personas cualificadas de una empresa y sus competencias, posición en el mercado, buena voluntad, reconocimiento, logros, patentes, contactos, soporte, colaboradores, liderazgo, base de clientes fieles, y la reputación son algunos de los activos intangibles clave que son difíciles de cuantificar. [3]

Siendo el capital intelectual uno de los principales activos, para lograr esa experiencia colectiva, cobran importancia el trabajo en equipo, la comunicación, la colaboración, el liderazgo y elementos que promueven el incremento de estos activos intangibles y que permiten aumentar el grado de conocimiento de cada integrante de un

equipo de trabajo. La tabla No. 1 muestra la categorización del grado de conocimiento de una persona en un dominio.

Tabla No. 1. Categorización del grado de conocimiento de una persona en un dominio, de Henao M, Calad M, Arango P., Fonnegra A., 2016.

Categoría	Descripción
	Aquel que es totalmente ignorante de un dominio
Novato	Quien ya ha tenido algún tipo de instrucción introductoria en un dominio
Aprendiz	Literalmente alguien que está en proceso de aprendizaje, convive con alguien que está en un nivel más alto en el dominio y lo asiste y acompaña
Competente	Persona que está en capacidad de ejecutar labores cotidianas sin supervisión directa. Es un trabajador experimentado pero sin alcances creativos
Experto	Es una persona competente, distinguida y brillante. Con conocimientos teóricos y experienciales en varios dominios. Altamente valorada por sus pares y colegas porque sus juicios son muy acertados y su desempeño muestra habilidad y destrezas fuera de lo común. Es creativo y puede manejar casos extremos y extraordinarios.
Maestro	Es el experto reconocido por sus colegas como el más experto; además calificado como guía y tutor para aquellos que están en una categoría menor en la escala, sus juicios son tomados para definir regulaciones y parámetros estándares.

Los equipos de trabajo efectivos que crean y usan información requieren la habilidad para: facilitar reuniones productivas, virtuales o presenciales y la comunicación entre la gente, Almacenar y obtener resultados de esas interacciones y; entender cuando los resultados del equipo y comunidad tienen significado para comunidades más amplias.[4]

En el aspecto de la comunicación se requiere:

1. Expresar ideas de forma clara escrita, oral o en una presentación.
2. Escuchar y evaluar las opiniones y la información que otros aporten.
3. Asignar los medios más efectivos para la comunicación en circunstancias particulares

Integrando el capital intelectual, las herramientas que permiten incrementarlo y el grado de conocimiento de las personas con la infraestructura, permite potenciar la administración de conocimiento, donde las tecnologías de información son un componente de la infraestructura.

Para la administración de conocimiento es esencial reconocer que la tecnología permite cambios en la cultura y operación de una organización y estos cambios son fundamentales, esto lo hacen posible las Tecnologías de Información que proporcionan ventajas muy importantes y permiten una relación simbiótica entre la administración de conocimiento y las organizaciones virtuales.

Cada época histórica tiene distintas formas organizacionales; de forma similar a la era industrial y su burocracia, la forma organizacional de la era de la información pueden ser las redes extendidas-limitadas. Básicamente una organización virtual es una red de trabajo asociada. El término “virtual” viene del latín *virtus* que significa

aprovechamiento, adelanto y define un atributo de una cosa, que no existe realmente, pero tiene la posibilidad de existir. [2]

Una organización virtual se basa en tecnologías de información que le permite utilizar recursos y competencias dispersos espacialmente, así como compartir información y conocimientos dentro de la red creada por la empresa y sus partes interesadas. El trabajo en una organización virtual se realiza fuera del marco del tiempo, y lugar, tomando la forma más común de trabajo remoto (es decir, el teletrabajo) como parte

de equipos de trabajo virtuales. El modo de acción está determinado por el propósito y los beneficios de la cooperación realizada, a menudo como parte de proyectos a corto plazo. Las organizaciones virtuales se caracterizan por una gran flexibilidad y capacidad para responder a los cambios. En el entorno turbulento. La gestión de la organización virtual es un proceso complejo, basado en menor medida que en las empresas tradicionales en la realización de las funciones clásicas de gestión y dependientes de muchos factores, tales como procesos de globalización, diferencias culturales, poder de negociación y alcance de los objetivos. [6]

El conocimiento, que es el activo intangible más importante de las organizaciones virtuales, es generado sobre la base de datos e información, y como resultado de un proceso de gestión eficaz, puede transformarse en sabiduría, entendida como la capacidad de uso del conocimiento en la práctica. La gestión del conocimiento está asociada con el capital intelectual de las organizaciones y determina su capacidad para aprender y responder con flexibilidad a los cambios en un entorno turbulento. Las organizaciones virtuales operan dentro del economía compartida, en la que las actividades relacionadas con la transferencia de información y el conocimiento, tanto dentro de las empresas como entre ellos y las partes interesadas, es de importancia particular. La creación, difusión y aplicación del conocimiento. El uso de métodos apropiados y herramientas tecnológicas debe ser una acción prioritaria para los gerentes de organizaciones virtuales que deseen lograr y mantener una ventaja competitiva. [6]

Se han planteado las fortalezas, ventajas y beneficios de la AC, sin embargo varios autores identifican algunas de las debilidades, cuando se implementa la gestión del conocimiento en las organizaciones, entre ellas [3]:

1. Los empleados no pueden encontrar el conocimiento crítico existente a tiempo. Ejemplo: "La solicitud: Se le pide a la compañía que ofrezca una cotización para un cliente importante. Cotejar lo necesario: información de los registros de la compañía o rastreando a sus propios consultores con la experiencia relevante se convierte en una tarea irrealizable en el marco de tiempo permitido. Cumple con la fecha límite, pero sus documentos de licitación están lejos de ser perfecto. La compañía pierde la oferta a un competidor".

2. Las lecciones se aprenden pero no se comparten. Ejemplo: "Usted nota que su oficina en Atlanta es trayendo mucho menos ingresos que su oficina en Boston, a pesar de que básicamente, hacen el mismo trabajo y atienden a una base de clientes idéntica. Las lecciones aprendidas y las mejores prácticas seguidas por los empleados de su oficina de Boston están siendo aprendida de nuevo por los de Atlanta. No hay ni un proceso suficiente ni la infraestructura necesaria que permita compartir o transferir las mejores prácticas a través de las dos oficinas. El intercambio de empleados durante algunos meses no ayudó, aunque la compañía pensó que si lo haría".

3. Su empresa no puede mantenerse al día con la competencia. Ejemplo: "Tu mayor

competidor parece para ganar nuevos clientes a un ritmo más rápido que tu empresa. También está perdiendo menos clientes que los que tú pierdes. Su empresa no parece estar aprendiendo de sus errores recientes, y sus competidores parecen estar aprendiendo tanto sobre sus errores como sobre nuevas oportunidades a un ritmo más rápido que tú".

Las Tecnologías de Información como herramientas de soporte para la Administración del Conocimiento

Los temas centrales que dominan el campo de administración del conocimiento son: Aprendizaje Organizacional (AO), Administración de Conocimiento y Tecnología. La Administración de Conocimiento ha llegado a ser una forma de capturar el "expertise" completo de una institución, dirigiéndolos a tecnologías como: bases de datos, interfaces web y documentos; conocimiento de la infraestructura para un acceso global del conocimiento de la organización, compartiendo conocimiento no sólo con la institución sino con clientes externos; una cultura de conocimiento de la institución, capturando el conocimiento tácito y la experiencia del personal, información de bibliotecas, centros de registro, unidades administrativas, operativas y del personal individual. [7]

Los especialistas en AO se enfocan hacia la investigación en tecnología por las instituciones para transferir la información y conocimiento y hacerlos disponibles a nivel organizacional. Los expertos en AO también dicen que las Tecnologías de Información nunca dirigen el conocimiento tácito el cual incluye no sólo las acciones, *expertise* e ideas del personal sino también sus valores y emociones. EL AO enfatiza que la eficiencia y efectividad de los trabajadores del conocimiento depende mayormente de cómo se comunican y colaboran en sus campos de trabajo y se presentan y exponen a su comunidad en la institución, así como fuera de ella.

Utilizamos el capital instrumental con las herramientas adecuadas, porque no es solo cuestión de la herramienta sino de cómo la utilizemos ya que los resultados pueden ser muy distintos. Incluso, podemos llegar a tener las mejores herramientas pero no necesariamente los mejores resultados si no logramos hacer uso efectivo de las mismas. El capital instrumental nos permite entonces potenciar la generación de valor que buscamos gestionando el conocimiento.

A continuación se presentan las tecnologías, conocidas como estrategias tecnológicas para apoyar la AC, pues facilitan la gestión de los activos de conocimiento. Algunas de ellas están orientadas a identificar los conocimientos explícitos que se tienen o que se necesitan [5];

1. Mapas de conocimiento
2. Páginas amarillas
3. Comunidades de expertos
4. Comunidades virtuales
5. Comunidades de práctica
6. Lecciones aprendidas
7. Cafés de conocimiento, k-café
8. Portales de conocimiento
9. Modelo de conocimiento en mapas conceptuales.

El interés de los expertos fue integrando al capital intelectual y el capital social (conocimiento tácito) con estas áreas para vender productos como sistemas de administración del conocimiento (SAC). Un sistema de administración de conocimiento consiste de los siguientes subsistemas [6]:

1. Búsqueda (mapas de conocimiento, redes)
2. Representación del conocimiento (gráficas de proceso, hipertexto),
3. Repositorio (archivos, bases de datos, sistemas de administración),
4. Colaboración (foros, correo administración de procesos),
5. Seguridad (Firma electrónica, cifrado),
6. Integración de recursos de información (sitios web, *datawarehouse*, sistemas de gestión documental).

Calidad en los servicios de Tecnologías de Información

Se define Calidad, como el grado en el que un conjunto de características inherentes cumple con los requisitos acordados con su cliente. La Calidad es sinónimo de eficiencia, confiabilidad, facilidad de uso, seguridad e integridad. Dentro de Tecnologías de Información, hay normas, estándares y mejores prácticas, donde se relaciona la gestión del conocimiento, ITIL es un marco de referencia que apoya la administración de los Servicios de TI, por documentar las mejores prácticas relacionadas con el suministro de servicios de calidad en la gestión de las Tecnologías de la Información. ITIL supone un enfoque para la gestión de los servicios TI basado en la Calidad de Servicio, concretándose ésta en los llamados Acuerdos de Calidad de Servicio (ACS) o Acuerdos de Nivel de Servicio (SLA). [8]

En ITIL la Gestión del Conocimiento es parte del proceso de transición del servicio y tiene como objetivo: recopilar, analizar, archivar y compartir conocimientos e información dentro de una organización. El propósito primordial de esta gestión es mejorar la eficiencia reduciendo la necesidad de redescubrir conocimientos. La Gestión del Conocimiento se encarga de establecer unos criterios de registro y de acometer labores periódicas de clasificación, evaluación y mejora de los datos disponibles. [9]

Una buena Gestión del Conocimiento ha de colaborar estrechamente con los procesos de las otras fases del Ciclo de Vida para documentar y analizar:

1. Los errores detectados y las soluciones aportadas en cada caso, principalmente desde la Gestión de Incidencias y Errores. De esta manera, puede confeccionarse un registro que recibe el nombre de KEDB y que ayuda a minimizar el tiempo de catalogación y solución de los mismos en el futuro. Asimismo, la Gestión de Problemas puede hacer un seguimiento del histórico de errores, establecer relaciones y determinar con mayor facilidad las causas de los mismos.

2. La Gestión de Cambios aportará documentación sobre las propuestas de cambio llegadas desde la fase de Mejora Continua del Servicio, tanto si han sido preaprobadas como si se han desechado.

3. La información relativa a las posibles consecuencias del error, que puede proporcionar al Centro de Servicios la posibilidad de anticiparse al cliente.

La Gestión del Conocimiento es la encargada, por último, de centralizar toda esta información en un repositorio denominado Sistema de Gestión del Conocimiento del Servicio (SKMS).

Otros marcos de referencia son COBIT, que está orientado la gestión y el gobierno de TI en las empresas e ISO 9001:2015 que es una norma de calidad aplicable a cualquier tipo de organización [10], en la relación entre ITIL, COBIT e ISO 9001, observamos que como un punto común tienen a la gestión de conocimiento (figura 1)

Fig.
1.



Elementos comunes de normas y marcos de referencia para servicios de TI: ISO, COBIT e ITIL. Se puede consultar en la página <https://www.gb-advisors.com/es/normas-y-estandares-internacionales/iso-cobit-e-itil-1-esp/>

Dentro de los elementos comunes de estos estándares marco de referencia podemos describir el proceso de la gestión que se requiere para encontrar la información como un elemento fundamental para la gestión de problemas [11]. La definición del problema es una habilidad analítica y requiere una definición sistemática de la información necesaria (Qué, quién, cómo, cuándo, dónde, por qué), se describen los elementos:

1. Iniciar la búsqueda – requiere un nivel básico de habilidades de búsqueda. Requiere juicio para decidir que fuentes son requeridas, una apreciación de las más apropiadas

2. Formular estrategias de búsqueda – requiere un conocimiento detallado de las fuentes más apropiadas y su estructura. tormenta de ideas, herramientas de búsqueda, palabras clave, conceptos, criterio de evaluación.

3. Uso de la información. El uso de información estructurada, la cual es utilizada de forma constante, requiere más que la habilidad de seguir instrucciones predeterminadas y competencia para utilizar las herramientas asociadas. El diseño en sí mismo es elemento importante de la eficiencia de operación y la habilidad de utilizarlos. Pero también es necesario habilidades para: evaluar la información encontrada, valorar su relevancia, calidad y decidir si es adecuada a las necesidades, Integrar la información de fuentes dispares, seleccionarla de forma efectiva, seleccionar formatos apropiados e interpretar la información

4. Creación de información. A menudo es producto de los procesos del negocio y los resultados podrían no ser reconocidos como información como documentos de todo

tipo que se envían a clientes, proveedores, autoridades o en forma de reportes. Para ello es necesario un entrenamiento efectivo entre cuyas acciones críticas se tienen: Decidir propósitos, identificar palabras clave, seleccionar formatos apropiados, definir estándares, crear un producto original (con herramientas adecuadas texto, audio, video, etc.), entender cuestiones legales y corporativas y proveer guías apropiadas, mapas, documentación.

5. Compartir información. En el manejo o administración del conocimiento un valor real es compartir ideas, hacer conexiones entre ellas, explorarlas y explotadas. Las herramientas clave son la comunicación y el trabajo en equipo.

6. Metodologías y prácticas utilizadas

En esta sección del documento se describen la metodología y prácticas utilizadas para desarrollar el "Modelo de gestión de conocimiento para la atención a usuarios y soporte de TI en la UADY".

Entrevistas con el personal

Se realizaron diversas entrevistas con personal clave en la administración de la RIUADY: responsables de las áreas operativas, con los encargados de la creación, recopilación y catalogación de la información, así como con el encargado de la implementación de la base de conocimientos para conocer el proceso que se siguió para implementarla. Se encontró que la implementación de la base de conocimientos utilizó gran cantidad de la documentación de procesos que se tenían almacenados durante varios años en carpetas físicas y archivos digitales, principalmente se contaba con información sobre los servicios en red.

Se encontró que se realizó un proceso para volver a catalogar los servicios principales de acuerdo a la experiencia del personal que brinda la atención a usuarios y el soporte técnico. Durante este proceso se eliminaron procedimientos obsoletos, se depuraron y optimizaron y se crearon nuevos con la experiencia obtenida.

Se pudo apreciar que el proceso de flujo de información y resolución de problemas se apegaba en gran medida a lo observado en la literatura sobre flujo de conocimiento. Se notó una gran disposición de parte del personal para el aprendizaje organizacional y el intercambio de información.

Recopilación de información para proponer el modelo de gestión de conocimiento

Se revisaron procesos, procedimientos, políticas, código de conducta, estructuras de trabajo, capacitación recibida y documentos de planeación estratégica de TI, se verificaron los permisos y características de seguridad que se maneja en la institución. Se decidió que la forma óptima para el manejo de información sería mediante un sistema de base de datos que soportara una interfaz vía Web y que opera en la nube.

En las visitas aleatorias realizadas, se observó la ejecución de la operación de los servicios de TI basada en los procedimientos y como se proporcionan los servicios a

los usuarios de la UADY. Se encontró que la información generada se ha ido almacenando de diferentes formas tanto escrita como digital apoyados en diversos formatos digitales predefinidos que siguen un estándar interno.

Se encontró que existe un autoaprendizaje del personal de acuerdo a un esquema que se les propuso en el año 2003 y que se presente en la figura 2

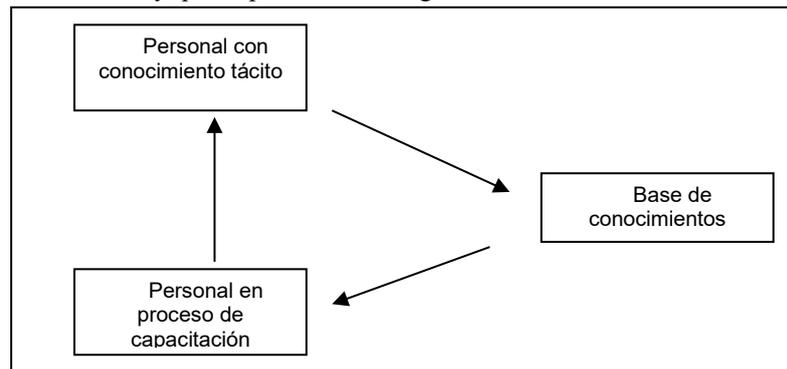


Fig. 2. Diagrama de autoaprendizaje del personal de TI utilizando las bases de conocimientos, Briceño Miguel, 2003.

También se observó un esquema incipiente de administración de conocimiento para proporcionar los servicios de red, que se implementó en el año 2005 y que se presenta en la figura 3:

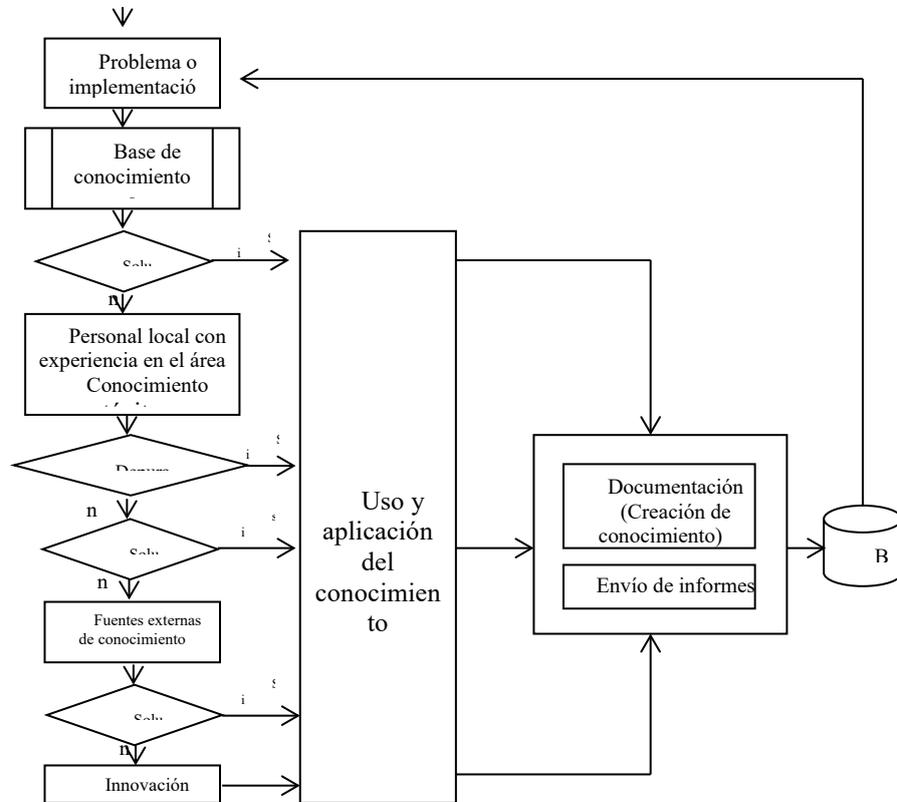


Fig. 3. Esquema de uso y aplicación del conocimiento de servicios en red, Briceño Miguel, 2003.

Implementación de las bases de conocimiento

La UADY optó por desarrollar un sistema propio, llamado Sistema de Atención a Usuarios (SAU), el cual tiene como objetivo ofrecer a los usuarios, un único punto para reportar o solicitar servicios de TI, asegurando así el correcto seguimiento y atención de sus requerimientos. Para la atención por parte de una mesa de servicios, el SAU cuenta con módulos que cubren cada paso del ciclo de atención incidentes definido por ITIL, incluyendo la Base de conocimientos, a través de la cual se actualiza la información técnica y procedimental, una imagen del SAU se muestra en la figura 4.



Fig. 4. Captura de una imagen del sistema de atención a usuarios (SAU), desarrollo propio de la Universidad Autónoma de Yucatán.

7. Resultados

La automatización de procesos es uno de los grandes cambios en la sociedad del conocimiento y el uso de Tecnologías de Información ha dado un nuevo giro al manejo del conocimiento organizacional. La Administración del Conocimiento por su parte está revolucionando los conceptos que hasta hace unos años se tenían en el manejo de la información [12]. Esta combinación ha permitido a instituciones como la UADY la creación de bases de conocimiento digitalizadas que permiten organizar y almacenar procedimientos de atención a usuarios, soporte técnico y provisión de servicios de TI a través de varios meses o años de investigación y acumulación de experiencias, así como aprovechar el conocimiento tácito y el conocimiento explícito del personal. Como resultado de la integración de los diversos elementos identificados en el estudio: capital intelectual, procesos de TI, sistema de atención a usuarios y el capital instrumental, se obtuvo el modelo de gestión de conocimiento para la atención a usuarios y soporte técnico, mostrado en la figura 5:

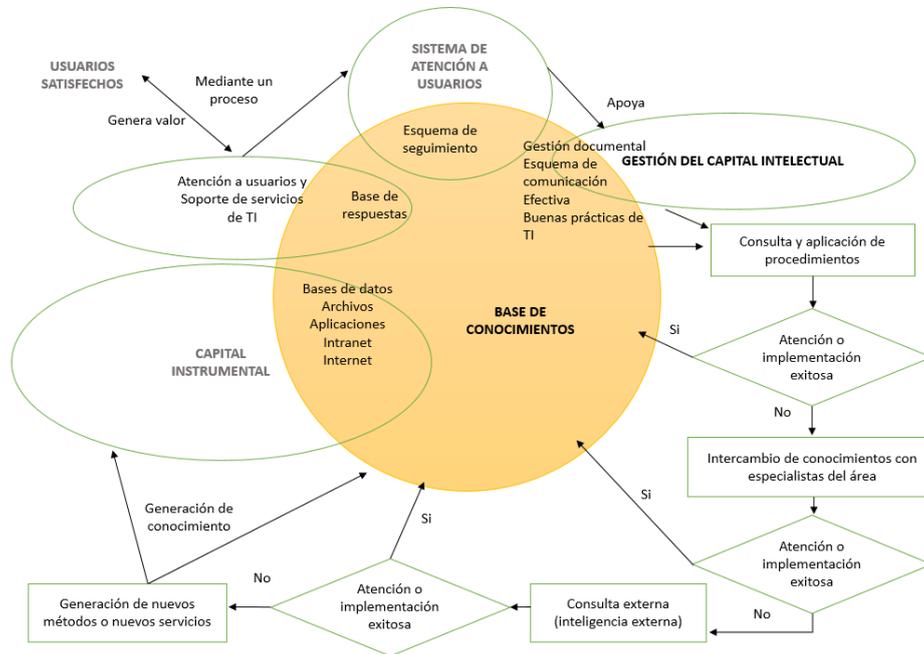


Fig. 5. Modelo de gestión de conocimiento para la atención a usuarios y soporte de servicios de TI en la Universidad Autónoma de Yucatán, el modelo es una propuesta propia y está basado en el Esquema de uso y aplicación del conocimiento de servicios en red de Briceño Miguel, 2003; y en el Modelo de gestión del capital humano en la RIUADY de Loeza David, 2006.

El modelo de gestión de conocimiento para la atención a usuarios y soporte técnico propuesto tiene elementos de ISO 9001:2015, ITILv3 y COBIT5, siguiendo un modelo de transferencia de conocimiento multinivel para la atención de los siguientes subprocesos: soporte a gestión de incidentes, registro y categorización de incidentes, resolución de incidentes por parte del soporte de primera, segunda y tercera línea, gestión de incidentes graves, monitorización y escalado de incidentes, cierre y evaluación de incidentes.

Con base en los procedimientos de Soporte y Provisión de Servicios de TI (P-SG-CATI-01, y P-SG-CATI-03) la mesa de servicios lleva a cabo la atención de requerimientos de los usuarios de la UADY. Estas actividades requieren análisis, validación, factibilidad y ejecución de las actividades técnicas necesarias para atender una solicitud o problemática específica por parte de la mesa de servicios. Para llevar a cabo las actividades técnicas se requiere de acceso a la Base de Conocimientos que contiene instructivos de implementación y resolución de problemáticas previamente documentadas y probadas en el entorno UADY, y en caso de que la el instructivo no exista, se escala con el especialista de más alto nivel quienes proporcionan la asesoría para realizar la atención y posteriormente solicita la documentación instructivo que es validado técnicamente por el responsable del área, así como por parte del área de gestión para formato, coherencia y pertinencia, lo anterior se resume e ilustra en la figura 6.



Fig. 6. Etapas para la integración del personal y evolución de su grado de conocimiento para llegar a ser Tutores técnicos quienes se vuelven expertos en un área de dominio de las tecnologías administradas en la Universidad Autónoma de Yucatán.

Estas herramientas también están permitiendo al personal de nuevo ingreso aprender sin necesidad de capacitaciones formales adicionales, optimizando con ello el tiempo que el personal con experiencia perdía en estas labores. También se ha reducido de forma considerable el tiempo perdido en soporte a usuarios para problemas comunes y “fácil” solución.

Como valor agregado a esta metodología el personal que la utiliza ha conseguido una superación personal y un autoaprendizaje en materia de Tecnologías de la Información y la comunicación que ellos mismos dudaban en alcanzar en tiempo tan corto, los tiempos de formación para conseguir el grado de especialista en un área fueron transformándose y disminuyendo de 2 años a y medio a dos años un mes pero logrando iniciar la formación como especialista en una año y medio, como se ilustra en la figura 7.

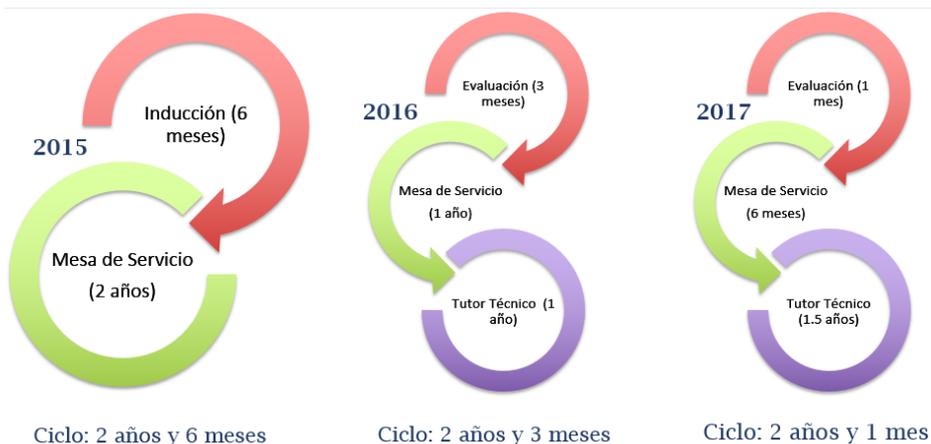


Figura 7. Evolución de las etapas y tiempos de incorporación del personal, para lograr acortar la transferencia de conocimientos especializados en un área de TI, incrementando el grado de conocimientos en un área de dominio de las Tecnologías de información administradas en la Universidad Autónoma de Yucatán.

Se obtiene que no es necesario hacer una mayor inversión para implementar una base de conocimiento que la que ya se tiene en cualquier Institución de Educación Superior. Más bien se necesita una buena labor de Gestión de conocimiento y Gestión de Tecnologías de Información, conocer de forma detallada los temas que se desean tratar, disponibilidad del personal y profesionales para compartir las lecciones aprendidas, conocimientos y experiencias, asesoría técnica para la implementación de los motores de búsqueda y una constante retroalimentación con deseos de aprender y enseñar.

Todavía nos deparan muchos cambios en la tecnología y el manejo de información y debemos estar preparados para ello y este tipo de manejo de información aplicando un modelo de gestión de conocimiento para la atención a usuarios y el soporte de TI.

Beneficios

Los beneficios que la Universidad Autónoma de Yucatán ha obtenido aplicando este modelo son:

1. Reducción de tiempos de integración del personal de nuevo ingreso ya que es posible capacitarse técnicamente sin necesidad explícita de capacitación.
2. Reducción del tiempo de soporte a usuarios para problemas comunes y servicios estandarizados, por ejemplo la atención de problemáticas con correo electrónico.
3. A pesar de la demanda exponencial y crecimiento de los servicios de TI, se ha logrado conservar los mismos niveles de atención y calidad de servicio, con el mismo personal.
4. Reducción de tiempos de implementación de nuevos servicios TI, mediante la aplicación de buenas prácticas basadas en procesos de ITIL y buen gobierno.
5. Se ha establecido un plan de capacitación especializada, basada en perfiles de las áreas técnicas.
6. Evaluación de habilidades y capacidades del personal con enfoque de 360 grados.
7. Disminución en el impacto de rotación del personal de la mesa de servicios.
8. Creación de bases de conocimiento dinámicas en un modelo colaborativo.

Conclusiones

El modelo de gestión de conocimiento en la UADY, ha demostrado ser una herramienta que podría ser aprovechada en otras áreas del conocimiento y por otras áreas de TI en Instituciones de Educación Superior.

Los profesionales de las Tecnologías de Información debemos darle un mayor impulso en nuestras instituciones a la creación y gestión de bases de conocimiento y demostrar las bondades que brinda. Pero definitivamente también es necesario abogar para que las Tecnologías de Información no se vean sólo como un gasto en

infraestructura y personal sino como una inversión que a corto, mediano y largo plazo puede dar excelentes resultados, siempre y cuando sean bien administradas y utilizadas. Se puede tener la última tecnología arrumbada y depreciando su valor o esa misma tecnología e incluso menor en capacidades, resolviendo una gran cantidad de problemas en materia de Tecnologías de Información.

Las Tecnologías de Información han dado un auge sin precedentes al manejo de información en cualquier área del conocimiento, de hecho, es casi imposible ahora, concebir un manejo de información sin utilizar herramientas informáticas. Estamos en un punto en el que es imprescindible aprovechar al máximo las TI disponibles en nuestras instituciones para abatir los rezagos en materia de aprovechamiento del conocimiento. Debe crearse una cultura que permita ver de manera natural el compartir e intercambiar información en cualquier área del conocimiento. [14]

Las demandas en el ecosistema de Internet deben abrirnos los ojos para ver que los profesionales de hoy que no se mantengan actualizados en sus conocimientos y su experiencia corren el riesgo de ser desplazados por una nueva generación de profesionales con habilidades muy diferentes a los que se tenían hace tan sólo 10 años.

Las perspectivas con el modelo de gestión de conocimiento para las áreas de TI de la UADY es que en sí mismo el modelo continúe evolucionando y madurando a otros niveles, abarcando a más personal de TI y alimentándose de su propia base de conocimientos [15].

Algo que resulta de suma importancia es que se debe aprovechar la mayor cantidad del conocimiento que nuestro propio personal sabe de forma tácita. Será necesario implementar nuevas formas de recolectar el conocimiento para evitar que se pierda en el retiro o fallecimiento de nuestros empleados o lo que podría ser peor que ese conocimiento pudiera filtrarse hacia fuera y utilizarse por otras organizaciones a su favor o en perjuicio de las nuestras.

Lecciones aprendidas en cuanto a buenas prácticas de TI

En esta sección se listan aquellas acciones que han dado buenos resultados en la Universidad Autónoma de Yucatán y que, de ser replicarlas bajo circunstancias o contextos similares, podrían dar los mismos o mejores resultados.

1. El modelo de gestión del conocimiento permite brindar soporte y aprovisionar los servicios de TI con los niveles requeridos con un número limitado de personas incluyendo a personal que realizar prácticas o servicio social, en el momento de la realización de este trabajo 13 personas de base y 9 becarios.
2. El modelo de gestión del conocimiento permite ofrecer a los estudiantes un esquema de formación que los prepara para los retos profesionales de su formación académica en diversas áreas de TI: telecomunicaciones, redes, servicios en la nube, plataformas Windows, seguridad informática y buenas prácticas de TI.
3. Para la correcta integración de un nuevo personal se requiere un periodo que permita su adaptación, así como la evaluación de sus capacidades, aptitudes y conocimiento tácito.
4. Para una correcta integración del personal se requiere un esquema de

inducción que permita al nuevo personal conocer los lineamientos, políticas de operación y buenas prácticas que se deben seguir fuera de las actividades técnicas.

Limitaciones y dificultades para adoptar el modelo

Sin duda, compartir las limitaciones o dificultades que se presentaron durante el proyecto, será de beneficio para otras instituciones interesadas en adoptar el modelo, tratando de evitarlas o tomándolas en consideración para prevenir problemas similares en otros proyectos bajo circunstancias o contextos parecidos.

1. Personal sindicalizado o con resistencia al cambio cuyos intereses son ajenos a su formación y/o desarrollo de las TI, enfocándose únicamente en el sueldo, problemas personales, actitudes negativas, vicios culturales y sociales. Lo cual imposibilita o impide la transferencia de conocimiento.
2. Existe personal que ha laborado en otras empresas e instituciones, de manera informal sin apego a procedimientos, que tienen dificultad para seguir instrucciones y procedimientos, afectando su rendimiento y la provisión de servicios de TI.
1. El presupuesto asignado para el pago a estudiantes (becarios) y personal por honorarios es limitado, lo cual genera mucha rotación de personal.

Agradecimientos

A mi Dios por la vida, la familia, la salud y el trabajo.

A mis hijos, Sofy, Juan Pablo y Samanta que siempre me apoyan y me demuestran su cariño y comprensión.

A mi esposo Miguel Briceño, que con su trabajo en la Especialización en Administración de Tecnología sentó las bases para ir consolidando el modelo de gestión de conocimiento que actualmente se utiliza en la Universidad Autónoma de Yucatán.

A mi amigo David Loeza que creyó en este proyecto, que continuó su integración y difusión, y que desde el cielo me sigue acompañando en las iniciativas y proyectos que emprendo.

A mis amigos y compañeros en esta aventura de la gestión del conocimiento, Juan Herrera, Carmita Denis e Israel Novelo.

Al personal y alumnos del área de TI de la Universidad Autónoma de Yucatán que todos los días realizan su trabajo en el área de TI con compromiso y pasión.

Referencias

1. Briceño Quijano Miguel, Las TIC como Soporte a la Gestión del Conocimiento. Yucatán, México. Facultad de Ingeniería Química; Universidad Autónoma de Yucatán, pp. 4, 40-41 (2003).
2. Malhotra, Yogesh . Knowledge Management and virtual organizations; Idea group; publishing; U.S.A. Asis, (2000)

3. Tiwana, Amrit. The knowledge management toolkit : Orchestrating it, strategy, and knowledge platforms / Amrit Tiwana. Edición 2nd ed. Pie de Imprenta Upper Saddle River, N.J. : Prentice Hall PTR, Descripción 550 p. : il. (2002)
4. Beazley, Hamilton, La continuidad del conocimiento en las empresas : cómo conservar el conocimiento y la productividad cuando los empleados se van / Hamilton Beazley, Jeremiah Boenisch, David Harden. Pie de Imprenta Bogotá : Norma, Descripción 375 p. (2003).
5. Henao M, Calad M, Arango P., Fonnegra A., Soluciones tecnológicas que apoyan la gestión del conocimiento, AD-MINISTER Universidad EAFIT Medellín Número 8 ene - jul (2006)
6. Lipinzka, Aneta, Knowledge management in a virtual organization Jagiellonian Journal of Management vol. 1, no. 1, p. 65–76 doi:10.4467/2450114XJJM.15.005.3812 www.ejournals.eu/jjm, (2015)
7. Kanti, Srikantiah T., Koenig, Michael E.D ; Knowledge management For the information professional; Asis; U.S.A.; (2000)
8. Moyano Fuentes J., ... {et. al] Gestión de la calidad en empresas tecnológicas : de TQM a ITIL /. Pie de Imprenta Paracuellos de Jarama, Madrid : Starbook, 2010. 254 p. : il., gráf. ; 24 cm.
9. Jan van Bon ... [et al.], Fundamentos de la gestión de servicios de TI : basada en ITIL®V3 / Edición 3a ed. Pie de Imprenta [Zaltbommel] : Van Haren Publishing, (2008)
10. ISACA. COBIT 5: Un Marco de Negocio para el Gobierno y la Gestión de las TI de la Empresa. Rolling Meadows: ISACA. (2012)
11. Sánchez Peña, Juan José, et al. "ITIL, COBIT and EFQM: Can They Work Together?" International Journal of Combinatorial Optimization Problems & Informatics, vol. 4, no. 1, Jan. 2013, pp. 54–64. EBSCOhost, search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=85332537&lang=es&site=ehost-live.
12. Successes and Failures of Knowledge Management 1st Edition; Jay Liebowitz eBook ISBN: 9780128053379, ISBN: 9780128051870, Morgan Kaufmann Published Date: 14th June 2016 Page Count: 240
13. Loeza Dorantes David Enrique: Evolución y perspectivas de la administración de conocimiento en la Red Integral de la Universidad Autónoma de Yucatán, Universidad Autónoma de Yucatán, p.47 (2006)
14. Skrzypek, Elżbieta. "Paradigmas de la Gestión del Conocimiento en las Condiciones de Cambio del Entorno". Documentos de Investigación de la Universidad de Economía de Wrocław / Trabajos Científicos de la Universidad de Economía en Wrocław, no. 419, enero. 2016, pp. 189-206. EBSCOhost, doi: 10.15611 / 04.2016 419.17.
15. Akhavan, Peyman, and Maryam Philsoophian. "How to Increase Knowledge Management Maturity Level? - An Empirical Study in a Non-Profit Organization." IUP Journal of Knowledge Management, vol. 16, no. 3, July 2018, pp. 44–53. EBSCOhost.
16. Jay Liebowitz, 10 Successes and Failures of Knowledge Management", Morgan Kaufmann Published, (2016)
17. Klaus North, Gita Kumta, Knowledge Management: Value Creation Through Organizational Learning, Springer; Edición: 2nd, (2018)

18. Patiño Claudia, Referencia de internet, consultada en la URL:
<https://gestionaconocimiento.wordpress.com/> el 2 de mayo de 2019.

QRPower@UNGS-Activa La transformación digital universitaria a través de soluciones simples con códigos QR

Analia Barberio, Ruy Arce, Gonzalo Ribera, Pablo Martínez,

^aDirección General de Sistemas y Tecnología de la Información, J.M. Gutiérrez 1150,
B1613 Los Polvorines, Buenos Aires, Argentina
abarberi@campus.ungs.edu.ar, rarce@campus.ungs.edu.ar, gribera@campus.ungs.edu.ar,
pmartine@campus.ungs.edu.ar

Resumen. En el año 2016 en la Universidad Nacional de General Sarmiento se implementó el proyecto UNGS Activa, en el marco del proyecto de transformación digital UNGS^{xt}, con el fin de centralizar y mejorar la experiencia de la comunidad universitaria en el uso de los servicios digitales a través de una app mobile. De todos los desarrollos que se fueron incorporando a la plataforma Activa, nos centraremos en este trabajo en una serie de soluciones donde el punto en común es la utilización de códigos QR (Quick Response code) como punto de unión entre el mundo real y el virtual, y que a través de distintas formas de implementación, generaron soluciones y mejoras a problemas y necesidades concretas.

Palabras Clave: QR, Códigos, Apps móviles, UNGS Activa, realidad, virtualidad, unión, mundo, solución, ágil.

Eje temático: 3) Transformando los servicios hacia la mejora de la experiencia del usuario (comunidad universitaria) y 7) Respondiendo a la universidad cambiante de la era digital mediante desarrollos rápidos y flexibles.

1. Introducción

La Universidad Nacional de General Sarmiento (UNGS) es una universidad ubicada en el segundo cordón del conurbano de la provincia de Buenos Aires. Desde su fundación en la década de 1990, la misma ha tenido un crecimiento constante tanto en su matrícula como en la variedad de su oferta académica. Este crecimiento puso a sus áreas de gestión en el desafío constante de crear y ofrecer soluciones que mejoren la experiencia cotidiana de la creciente comunidad universitaria.

Entre estas áreas, la Dirección General de Sistemas y Tecnologías de la Información (DGSyTI), como responsable de la infraestructura tecnológica de toda la Universidad, se vio en la necesidad de generar soluciones digitales ágiles, que fueran sencillas de implementar y simples de desarrollar, y que sirvieran para acompañar la enorme transformación digital que fue evolucionando exponencialmente en los últimos años.

En este contexto, la DGSyTI gestionó y desarrollo el proyecto UNGS^{xt} con miras a acompañar la creciente y demandante metamorfosis digital universitaria. Dentro de este proyecto de amplio alcance, se gestó la plataforma UNGS Activa, destinada a montar una estructura ágil que permita integrar distintos servicios ya desarrollados, y además posibilite incorporar nuevas soluciones con desarrollos ágiles.

1.1 La App UNGS Activa

El proyecto UNGS^{xt} es muy extenso y no viene al caso explicitarlo en este trabajo, pero a grandes rasgos es un gran organizador, que ayudó a ordenar distintas entidades digitales en zonas con distintos fines. Por ejemplo podemos encontrar zonas de: personas, documentos, aplicaciones, seguridad, calidad, etc. Cada zona concentra y delimita las posibles soluciones al problema que enmarca.

A raíz de este proyecto, surgió en primer lugar un portal de servicios con el fin de agrupar bajo un mismo ingreso distintas soluciones que ya se encontraban en línea pero con accesos autónomos en cada una. Sobre este primer portal de acceso unificado, se desarrolló en paralelo la App "UNGS Activa", la cual contiene en sí misma los mismos accesos que contiene el portal web, pero que además extiende las posibilidades de solución al tener las ventajas propias de una aplicación móvil.

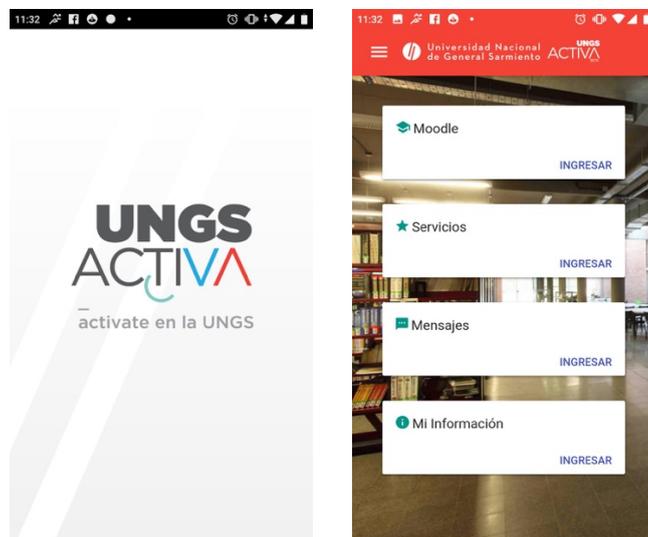


Fig. 1 y 2. Splash y Vista de la interfaz principal de la app UNGS Activa.

En el presente trabajo, mostraremos puntualmente dos casos de soluciones implementadas en UNGS Activa, que tienen en común la utilización de códigos QR, tanto la generación como la lectura desde la misma plataforma. En este sentido, intentaremos demostrar como la generación dinámica y la lectura de códigos puede ser una solución sencilla, eficaz y económica para la transformación de los servicios digitales.

1.2 Arquitectura y tecnología

La solución completa consta de un backend desarrollado en base a microservicios con APIs REST, un frontend web y una app móvil para dispositivos con Android y con iOS. Todas nuestras soluciones, tanto en backend como en frontend están desarrolladas en entorno MEAN +. PGSQL, esto es MongoDB + Express + AngularJS + NodeJS + PostgreSQL como base de datos principal. A las bases de datos también se suma REDIS como base de administración de sesiones de usuario.

A todo esto, se suma la utilización de Nuxeo como framework de soporte y administración documental, y en concordancia con este, se decidió también utilizar el estándar CMIS como puente entre nuestros desarrollos y dicho ECM. Por último, como módulo de componentes para los distintos frontends adoptamos la suite de componentes de Angular Material, inspirado en el concepto de Material Design de Google.

Como administradores de paquetes y librerías para organizarnos usamos NPM, Yarn y GULP como gestor de tareas para la puesta en producción de los distintos módulos. Por último, para hacer el seguimiento del avance de todo el código del proyecto y la interacción del equipo se implementó un servidor con GitLab, y para la comunicación interna y canales de notificación de novedades utilizamos Grupos y Bots de Telegram.

1.2 Los actores y las soluciones

Queremos en este punto explicitar brevemente las categorías sobre las cuales se desarrollará el presente trabajo. Es necesario en primer lugar delimitar que cada vez que se realiza la lectura de un código QR, se está estableciendo una relación entre un emisor pasivo de información (el código) y un lector activo (el dispositivo de lectura), el cual sensa activamente en todo momento en búsqueda del patrón conocido.

Estas dos posiciones de la relación, lo pasivo y lo activo, el código y su lectura, irán tomando distintos lugares en las soluciones propuestas y podrán estar identificados con roles y perfiles de usuarios o ser autónomos. Esto conforma otra categoría la cual entra en juego en la relación entre el código y su lectura, de acuerdo al actor que contenga la habilidad de generar la novedad o de leerla e impactar en el sistema.

Como ya dijimos, dadas las características de este trabajo solamente nos vamos a enfocar en dos soluciones distintas que tienen en común el reemplazo de circuitos administrativos centralizados en papel a circuitos digitales descentralizados. Estas soluciones, implementadas en distintos niveles en la plataforma Activa son en primer lugar el módulo de cupones de descuento, y en segundo lugar el proyecto *Registro de Actividades* en sus modalidad Docente (RAD).

2. Cuponera digital

La Universidad Nacional de General Sarmiento es pública y gratuita, y su comunidad estudiantil se compone principalmente de jóvenes de la zona noroeste del conurbano bonaerense, muchos de ellos de bajos recursos y que dentro de sus familias son la primera generación que puede acceder a educación universitaria. En este sentido, la UNGS siempre ha otorgado distintos beneficios para hacer más accesible la educación universitaria.

Desde la Dirección de Bienestar Universitario se impulsaron distintas acciones para ayudar a los estudiantes a continuar con sus estudios e intentar minimizar la deserción. Entre ellas se encuentran la bolsa de trabajo, las actividades de educación física y recreación, el programa de cuidado y prevención de la salud, el servicio de transporte circular gratuito y los cupones de descuento en el bar comedor universitario y en la librería del campus. Nos vamos a enfocar aquí principalmente en los cupones de descuento del bar, ya que los mismos no son ocasionales, sino que todos los días gracias a estos cupones los estudiantes activos pueden acceder a un menú estudiantil

Sobre este último punto de los cupones se basa la primera solución que traemos a este trabajo. Antes de la implementación de UNGS Activa, los cupones se entregaban en un papel impreso desde la ventanilla de atención del área de Bienestar Universitario. Esto implicaba que todos los días se imprimieran gran cantidad de cupones, que además al momento de la entrega se hiciera una verificación manual sobre el sistema de estudiantes para validar su situación académica (para acceder a los cupones es necesario ser estudiante regular activo y estar inscripto a materias). Esto ocasionaba largas colas de espera para retirar el cupón, múltiples controles en múltiples puntos (oficina de bienestar – bar), gasto excesivo en papel, etc.

2.1 La solución digital

Desde la Dirección General de Sistemas y Tecnologías de la Información (DGSyTI) nos propusimos desde la implementación del proyecto UNGS Activa desarrollar una solución que permitiera agilizar todo el circuito de los cupones de descuento, tanto desde su emisión como así también en la recepción en los distintos puntos donde se hace efectivo el descuento. Así se avanzó hacia la incorporación de un módulo de cupones digitales dentro de la misma aplicación móvil. Este módulo sería el primero en implementarse como solución dentro de la App además del acceso a servicios que brindaba la app por defecto.





Fig. 3 y 4. Vistas de acceso a la cuponera del menú estudiantil y al cupón diario.

Dentro de este módulo, los estudiantes que estén activos, es decir, aquellos que estén cursando materias en el cuatrimestre en curso, acceden desde la interfaz a un código QR que se genera en base a un hash creado en el momento de la solicitud del cupón. El cupón creado, genera en la base de datos centralizada de Activa, un registro con el vínculo entre el hash y los datos del estudiante que lo generó. Una vez registrado, hasta no ser canjeado el cupón permanecerá accesible desde la app durante el día completo. El estudiante que lo quiera utilizar simplemente debe dirigirse a la caja del bar o de la librería y mostrar su cupón de descuento.

Además del perfil del estudiante, el cual se valida contra el sistema de gestión académica SIU-Guaraní, y se asigna automáticamente a cualquier estudiante activo que se registra en la app, se creó también dentro de la misma un perfil de empleado de caja. Este perfil es asignado manualmente desde la DGSyTI, y permite a aquellos que lo tengan acceder al módulo de lectura de cupones. Este lector lo utilizan los empleados desde sus propios smartphones y les permite canjear códigos y verificar si los mismos son válidos o no. Un cupón QR una vez canjeado ya no puede utilizarse más.

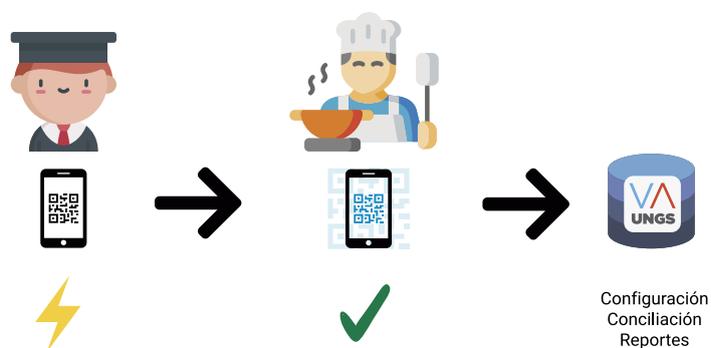


Fig. 5. Primero, el estudiante genera el QR desde su dispositivo y se lo muestra al cajero del bar, este, desde un dispositivo móvil con un perfil de lectura lo lee, y el momento del canje queda registrado en la base para posteriores conciliaciones de datos entre la Universidad y el Bar universitario.

2.2 La experiencia.

La app UNGS Activa fue desarrollada con IONIC 3 y liberada en Google Play Y App Store a mediados del año 2017. Desde su liberación se mantiene instalada en promedio en 3000 dispositivos activos (*) con android y unos 300 con ios (**). Es necesario aclarar en este punto, que para que los estudiantes accedan a su cupon, no es necesario tener la app instalada, ya que aquel que por algun motivo no pueda instalarla en su dispositivo el cupon tambien es accesible desde la plataforma activa publicada en activa.ungs.edu.ar y la cual es accesible via browser. No asi con el perfil de empleado de caja, el cual sí funciona unicamente dentro del formato app.



Fig. 6. Vista del QR del cupón que se presenta en la caja del bar universitario.

Además de la alternativa web, también se dejó una opción de una alternativa completamente analógica, donde aquel estudiante que no tenga acceso a ninguna de las opciones digitales puede pasar a retirar un cupon impreso en la oficina de Bienestar. Esta última alternativa que se mantuvo por previsión ante excepciones, podemos afirmar hoy que desde que fue liberada la versión digital no fue utilizada nunca. El alcance del beneficio se cuatuplicó en cuanto a la cantidad de estudiantes que acceden a él diariamente, en comparación a los registros en que se entregaban manualmente.



Fig. 7 y 8. Estudiantes utilizando la cuponera digital

En conclusión, esta primer herramienta implementada dentro de la plataforma Activa, es una muestra cabal de la transformación digital de un proceso sencillo dentro de la comunidad universitaria. Gracias a esta transformación además, se logró no solo agilizar el proceso, sino extenderlo aún más a estudiantes que no podían utilizarlo por cuestiones de tiempo o desconocimiento. Por último, un efecto colateral gratificante y no esperado, fue el reconocimiento del área de Bienestar Universitario, de que la herramienta no solo ayudó a optimizar y agilizar un proceso, sino que colaboró a que el estudiante no se sienta estigmatizado de alguna manera al tener que ir a buscar un descuento especial, dado que, al tenerlo ahora dentro de su app y su perfil con el tiempo lo han tomado como un derecho propio y para todos, y no como una necesidad excepcional a la que recurrían algunos a demanda y haciendo una cola.

(*) Fuente: Google Play Developer Console

(**) Fuente: Apple Developer

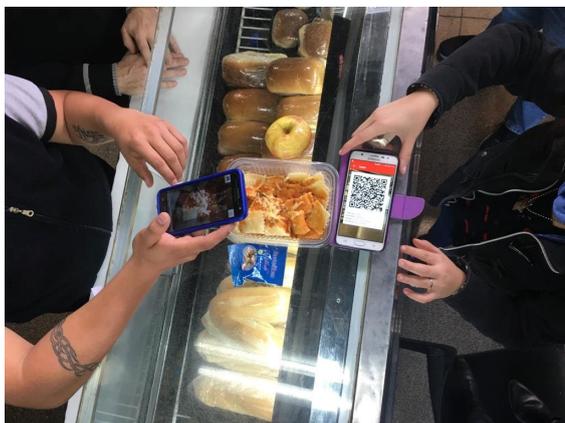


Fig. 9. Estudiante canjeando un cupón.

3. Registros de Actividad

En este segundo ejemplo de soluciones digitales implementadas donde el protagonista es el código QR, vamos a describir el módulo de Registro de actividades. El Registro de Actividad Docente (RAD) es un dispositivo mediante el cual los docentes que dictan clase, registran el tema que dictan en cada clase y con esto dan su presente y ocasionalmente alguna novedad (dictado fuera del aula, eventos especiales, avisos de ausencia con antelación, etc).

Antes de la implementación de la solución que vamos a describir, el RAD se llevaba en papel. La oficina de bedelia, todos los días imprimía el equivalente a un libro de páginas en blanco que iban a ser completadas por los docentes. Dicho libro de registro se ponía a disposición en la sala de profesores y los docentes pasaban por allí a registrar manualmente sus clases y sus novedades. Aunque el libro era dividido en varias partes para permitir el registro simultáneo, casi todos los días, en los momentos previos a los

turnos de mayor concurrencia era común ver un gran congestionamiento de docentes en las salas de profesores intentando registrar su actividad y dar su presente.

Además del registro manual que quedaba anotado en el libro diario, era necesario que toda esta información se vuelque en el sistema de gestión académica SIU-Guarani para poder luego cruzar información y obtener reportes sumariados. Esta carga no podía hacerse de otra manera que no fuera manual, y así era como varios recursos de la oficina de bedelia se veían afectados fulltime a pasar lo registrado a mano dentro del sistema, lo cual también afectaba en la calidad de la atención a los estudiantes y en la agilidad de otros procesos y trámites dependientes de la oficina.

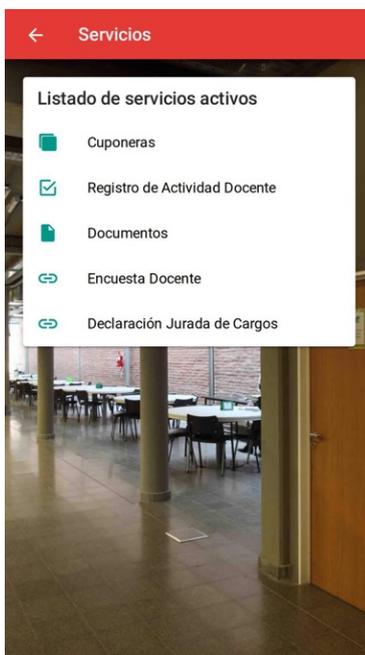


Fig. 10. UNGS Activa App: Vista de acceso al módulo de Registro de Actividad Docente.

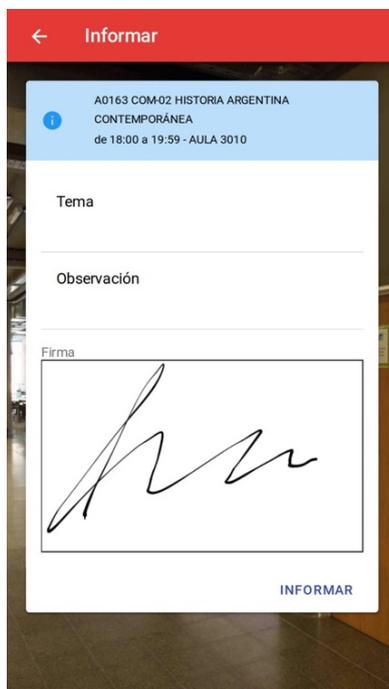


Fig. 11. Vista del listado diario de materias a informar.

3.1 La solución digital

A finales de 2018, y en buena parte gracias al significado y empuje que ocasionó el trabajo de actualización y migración del sistema de gestión académica SIU-Guaraní, nos pusimos manos a la obra en el desarrollo de una solución digital para la transformación del viejo circuito del RAD. La meta era realizar una aplicación, que de manera sencilla permita una descentralización de la carga en el sistema, que afectara la mínima cantidad de recursos humanos y técnicos. La solución además debería resolver el registro de la presencialidad del personal de manera no invasiva, es decir, sin utilizar una captura de datos biométricos.

Se decidió entonces desarrollar un sistema de carga descentralizada, donde los docentes informarían sus novedades a través de un formulario accesible desde la app UNGS Activa. En dicho formulario podría cargarse el tema de la clase junto a alguna observación (opcional), y también, la firma holográfica manual del docente. Una vez cargada y confirmada la información en el formulario, al confirmarlo es generado un hash único por docente, fecha y hora, el cual se muestra a través de un código QR en la app.



The image shows a mobile application interface for reporting an activity. At the top, there is a red header with a back arrow and the word "Informar". Below this, a blue box contains course information: "A0163 COM-02 HISTORIA ARGENTINA CONTEMPORÁNEA" and "de 18:00 a 19:59 - AULA 3010". The form has three main sections: "Tema" with a text input field, "Observación" with a larger text input field, and "Firma" with a signature box containing a handwritten signature. At the bottom right of the form, there is a blue button labeled "INFORMAR".

Fig. 12. Vista del formulario de registro de la actividad con firma manual holográfica.

En este caso, al contrario de la solución de la cuponera, la lectura de los QR debería hacerse en lo posible sin intervención de ningún actor para no poner recursos de bedelia solo a leer QRs de los docentes. Por tal motivo se decidió instalar lectores de QR autónomos, desarrollados con una Raspberry PI Zero junto a una cámara web, que toman el hash provisto por el código QR y se comunican con una api REST que hace el POST con la novedad al servidor. Toda la información capturada de las novedades de los docentes día a día, es registrada en nuestro Sistema Único Documental (SUDOCU) dentro de un documento del tipo "Registro de Actividad Docente" diseñado para tal fin. Este documento, una vez finalizado el día, es revisado desde la oficina de Bedelia y si los datos son correctos y no hay nada que agregar se firma digitalmente y el acto de la firma junto a los metadatos del documento quedan registrados en blockchain para evitar la alteración de esos datos en el futuro.



Fig. 12. Una vez informada la novedad se genera el código QR que el docente deberá presentar en los puestos de lectura automáticos.

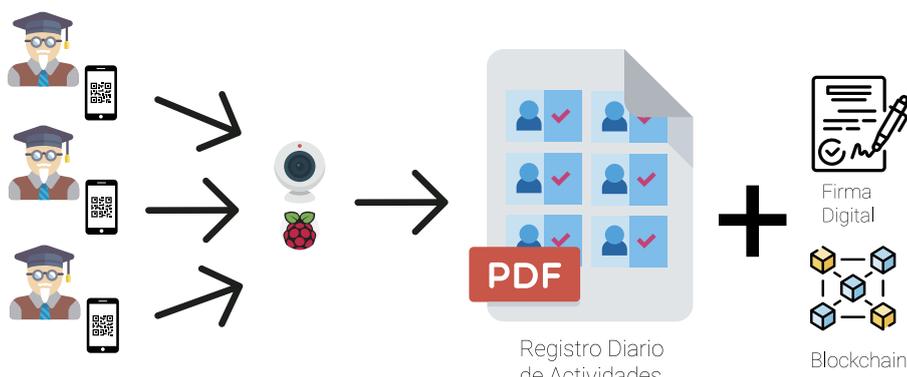


Fig. 13. Circuito del registro de actividad: Los docentes generan su QR informando la novedad en sus dispositivos móviles, luego pasan el QR por el lector Raspberry ubicado en la sala de profesores. Luego, todas las novedades completan un documento de registro diario que al finalizar el día es firmado por una autoridad de Secretaría Académica y registrado en la Red Federal de Blockchain de Argentina.

3.2 La experiencia

Si bien el propósito de este módulo, como se aclaró mas arriba, no es el de ejercer un control fuerte sobre el presentismo y el ausentismo, sino tan solo ser un registro de novedades utilizando una tecnología no invasiva. De todas maneras la primera implementación obtuvo una fuerte resistencia de algunos sectores docentes, sobre todo

9ª Conferencia de Directores de Tecnología de Información, TICAL2019
 "El genoma estudiantil y la metamorfosis digital universitaria"
 Cancún, México, 2 - 4 de septiembre de 2019

de las franjas etarias más altas, y de docentes que no acostumbran a utilizar la tecnología en sus acciones cotidianas.

Tipo y N°	Descripción	Estado	Origen	Fecha
<input checked="" type="checkbox"/> N°2 #011-076022	Registro de Actividad Docente - Fecha 02/05/2019	Registrado	Usuarior: Ruy Sebastian Ace Área: Dirección General de Sistemas y Tecnologías de la Información	02/05/2019 11:58
<input type="checkbox"/> N°1 #011-076024	Registro de Actividad Docente - Fecha 28/03/2019	Registrado	Usuarior: Ruy Sebastian Ace Área: Dirección General de Sistemas y Tecnologías de la Información	17/04/2019 17:10
<input type="checkbox"/> Convenio N° 1935 #001-076022	convenio de pruebas	Registrado	Usuarior: Ruy Sebastian Ace Área: Dirección General de Sistemas y Tecnologías de la Información	28/03/2019 11:28

Fig. 14. Los datos capturados y recolectados por los lectores de QR generan un documento diario de Registro de Actividad Docente en el módulo de Gestión Documental del sistema SUDOCU.

Esto hizo que el plan de implementación, si bien en un primer momento se había planificado como obligatorio desde principios de 2019, haya pasado a ser de uso optativo por el momento, y ser tomado como un plan piloto de prueba hasta tanto se considere su uso como obligatorio. Si bien se habían planteado distintas estrategias de registro alternativo para las excepciones en las que no se pueda utilizar un celular, en algunos casos se nota la dificultad de alejarse del registro y la firma en papel.

Título y Fecha	Actividad	Usuario	Acción
Instituto del Desarrollo Humano			
AM04 CDM-01 ALGEBRA II de 14:00 a 17:59 - AULA 7143	Casarrubá, Efra González, Narciso Blanco Severo, Gastón Alejandro	Asistir con éxito	✓
AM176 CDM-01 ANTRPOLOGÍA FILOSÓFICA de 10:00 a 14:59 - AULA 7071	Suazo, Gisela Patricia	Informar	
AM04 CDM-01 ARTES VISUALES I de 12:00 a 13:59 - AULA 7077	Kotickidky, Daniela	Informar	
AM04 CDM-01 CONCEPCIONES DE LA COMUNICACIÓN de 18:00 a 22:59 - AULA 7283 / 7285	Vázquez, Cecilia López, Adrián Lucas	Informar	
AM05 CDM-01 DERECHOS ECONÓMICOS, SOCIALES, CULTURALES Y AMBIENTALES de 10:00 a 13:59 - AULA 7104 / 7176	Cervera, Verónica Lucía	Informar	
AM113 CDM-01 ENSEÑAR EN LA ESCUELA SECUNDARIA Y EN EL NIVEL SUPERIOR de 09:00 a 12:59 - AULA 7076	Kirovsky, Sranslo Esmeralda Chavezsky, María Magdalena	Informar	
AM113 CDM-01 ENSEÑAR EN LA ESCUELA SECUNDARIA Y EN EL NIVEL SUPERIOR de 09:00 a 12:59 - AULA 7072	Kirovsky, Sranslo Esmeralda Chavezsky, María Magdalena	Informar	

Fig. 15. Además de los datos recolectados por las capturas de QR, los usuarios de la oficina de Bedelia pueden revisar lo informado o informar manualmente otras novedades si es necesario.

Los argumentos utilizados para ir en contra son variados. Desde el control o el no acceso a la tecnología necesaria, como así también falta de capacitación y problemas de conectividad. No obstante estas dificultades que se suscitaron a priori, la prueba piloto continua y con los registros realizados en este contexto, la oficina de Bedelia asegura ya haber reducido en buena medida la carga manual de novedades, y se espera que una vez limadas las asperezas y mejorado el circuito se pueda implementar como obligatorio a partir del segundo cuatrimestre de 2019.

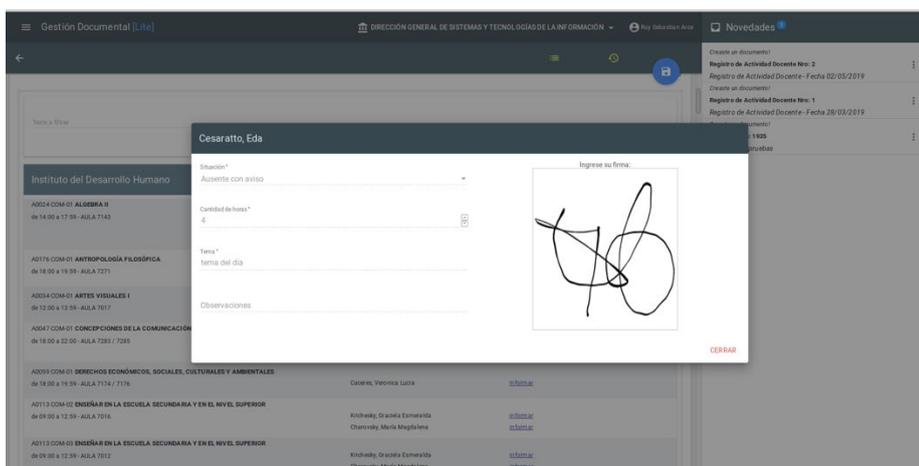


Fig. 15. Detalle de la novedad capturada desde la app del docente, vista en el sistema de gestión documental.

4. Conclusión

Hemos visto en el presente trabajo como ciertos procesos que tradicionalmente se gestionaban con papel y con una gestión centralizada, es posible, gracias a la generación al vuelo de códigos QR en una app móvil, descentralizar la generación de la información y así poder agilizar los circuitos. Continuamos actualmente trabajando en mejorar los procesos ya implementados como los que presentamos aquí, e incorporando nuevos desarrollos basados en la misma teoría y también extendiendo la misma.

Se avanza actualmente para que el Registro de Actividades no solo sea de las novedades docentes, sino que el mismo módulo de registro permita interactuar a un docente como concentrador de información y pueda capturar novedades de estudiantes. De esta manera un docente podría tomar asistencia digitalmente, no solo en el aula sino en clases que se llevan adelante fuera del contexto áulico.

Vimos entonces, como el código QR, si bien no es una herramienta que pueda servir como control primario, es sumamente versátil para implementar soluciones de registro simple, donde intervienen múltiples actores entre los que es necesario intercambiar información. El QR, como punto pasivo de sujeción de información, puede moverse de un actor a otro, o inclusive a un punto fijo del espacio, desde el cual es posible

interactuar a través de una app y ser a la vez el sensor activo que acceder a la información en ese punto del espacio-tiempo donde se encuentre.

Agradecimientos

A todo el pequeño gran equipo de la Dirección General de Sistemas y Tecnologías de la Información por acompañarnos a crecer y a la gestión Diker-Bonaldi por dejarnos ser.

Tablero de control universitario basado en el modelo de acreditación institucional ecuatoriano

Xavier Palacios^a, Ramiro Mosquera^b

^a Departamento de Sistemas, Dirección de Sistemas, Universidad Internacional del Ecuador, Av. Simón Bolívar y Jorge Fernández S/N, Quito-Ecuador
xpalacios@uide.edu.ec

^b Departamento de Sistemas, Area de Business Intelligence, Universidad Internacional del Ecuador, Av. Simón Bolívar y Jorge Fernández S/N, Quito-Ecuador
rmosquera@uide.edu.ec

Resumen. En este trabajo se explica el proceso que se siguió para construir el tablero de control de mando universitario (dashboard universitario), basado en el modelo de acreditación institucional ecuatoriano del año 2015 (vigente hasta la fecha), implementado por el CEAACES (Consejo de Evaluación, Acreditación y Aseguramiento de la Calidad de la Educación Superior) hoy llamado CACES (Consejo de Aseguramiento de la Calidad de la Educación Superior). Se hace referencia también a las bondades de la herramienta implementada y su impacto en la toma de decisiones de las autoridades universitarias. Este dashboard universitario fue construido con el conjunto de herramientas de análisis empresarial de Power BI de Microsoft y arranca con la obtención de la información del Data Warehouse(DW) que la institución tenía desarrollado años atrás, para generar un tablero de control con la utilización del método de semaforización con colores (verde, amarillo, anaranjado, rojo) que indican el estado de la institución dentro de todos los ámbitos del modelo de acreditación.

Palabras Clave: Business Intelligence (BI), tableros de control (dashboard), modelo de acreditación, data warehouse (DW)

Eje temático: Transformado los servicios hacia la mejora de la experiencia de los usuarios (comunidad universitaria)

1. Introducción

El desarrollo de las tecnologías de información y comunicación, han permitido y seguirán permitiendo un extraordinario desarrollo de la actividad humana. Tecnologías como: la inteligencia artificial (AI), el internet de las cosas (IoT), la analítica de datos, el blockchain, big data, entre otras, actualmente son los impulsores tecnológicos para esta nueva etapa de la transformación digital. Varias de estas tecnologías son parte del área de la inteligencia de negocios (BI), "término utilizado tanto por los distribuidores de hardware y software como por los consultores de tecnología de la información para describir la infraestructura para almacenar, integrar, crear informes y analizar los datos que provienen del entorno de negocios, incluyendo Big Data" (Laudon&Laudon, 2016). Como parte de estas herramientas de BI se tienen los tableros de control o dashboards, que pretenden dar una visión completa del estado actual de cualquier organización y brindar alertas para facilitar la toma oportuna de decisiones a nivel de la alta gerencia. En este trabajo lo que se presenta es justamente el proceso para llegar a construir un tablero de control "universitario", a través del cual las autoridades o los usuarios autorizados, tienen en una sola pantalla la visión completa del desempeño de la institución basado en un modelo de acreditación que toma en cuenta todos los ámbitos del quehacer universitario. Este trabajo fue desarrollado para la Universidad Internacional del Ecuador que tenía información dispersa en distintas bases de datos y archivos planos, lo que significaba un problema cuando se requerían reportes consolidados, a nivel institucional. El Departamento de Sistemas, ha venido trabajando arduamente desde hace varios años atrás, conjuntamente con la Dirección de Acreditación y el Vicerrectorado Académico para ir consolidando la información requerida para la generación de reportes útiles a la toma de decisiones. Luego de realizar esta consolidación con el Data Warehouse institucional, se generaron gran cantidad de reportes muy útiles, pero que no daban una visión más global de la Universidad. Para conseguir este objetivo, se tomó como referencia el Modelo de Acreditación Institucional promovido por el organismo de acreditación de las instituciones de educación superior del Ecuador. A continuación, se describe el método empleado para llegar a obtener esta solución que mejoró la experiencia de los usuarios que toman las decisiones estratégicas de la institución.

2. Desarrollo

2.1 Modelo de Acreditación

El Modelo de Acreditación propuesto por el CEAACES hoy CACES es la base para la realización de este tablero de control universitario. El modelo pretende medir el desempeño de las instituciones de educación superior con el firme propósito de mejorar la calidad del sistema de educación superior ecuatoriano. "Se entiende la calidad de las universidades y escuelas politécnicas como el grado en el que, de conformidad con su misión, enmarcada en los fines y funciones del sistema de educación superior

ecuatoriano, estas alcanzan los objetivos de docencia, investigación y vinculación con la sociedad, a través de la ejecución de procesos que observan los principios del sistema y buscan el mejoramiento permanente" (CEAACES, 2017). El modelo evalúa seis criterios: organización, academia, investigación, vinculación, recursos e infraestructura y estudiantes. Cada criterio tiene asociados subcriterios, los cuales son evaluados a través de magnitudes específicas llamadas "indicadores"(KPI's) que pueden ser cuantitativos o cualitativos. "En el caso de los indicadores cuantitativos, los estándares están implícitos en las funciones matemáticas que determinan el desempeño de la institución en una escala de 0 a 1, donde 0 representa un cumplimiento nulo del estándar y 1 corresponde al cumplimiento totalmente satisfactorio del mismo" (CEAACES, 2017). Para ilustrar de mejor manera lo indicado, se va tomar como ejemplo el criterio *Academia*, el subcriterio: *Posgrado* y el indicador: *Doctores a TC*. El estándar de calidad mide el porcentaje de Doctores (PHD) a tiempo completo en relación al número total de docentes a tiempo completo. La fórmula utilizada es:

$$DTC = (\# \text{ de profesores PHD a tiempo completo}) / (0.60 * \# \text{ de docentes a tiempo completo})$$

El valor entre 0 y 1, obtenido por la institución evaluada, depende de la función de utilidad (ver Fig 1). Como se puede observar la puntuación más alta (1) se obtiene con resultados de la fórmula superiores a 55 y valores menores o iguales a 10 no puntúan.

Utilidad:

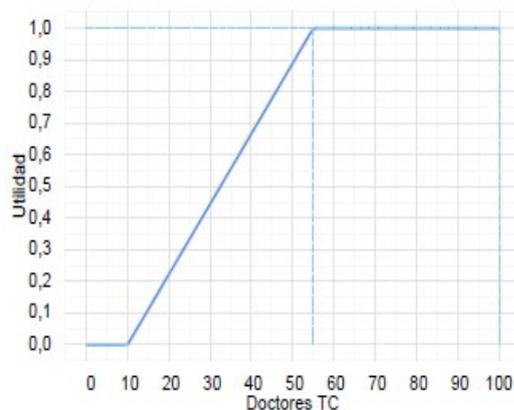


Fig 1: Función de utilidad de Doctores TC

Para el cálculo de estos indicadores los datos son obtenidos del data warehouse en base al proceso indicado en la sección 2.2. Para los indicadores cualitativos se consideran 3 niveles de evaluación: Satisfactorio (1), medianamente satisfactorio (0.5) y deficiente (0). Estos indicadores debido a su naturaleza son ingresados manualmente. Cada uno de los valores obtenidos en los indicadores, tienen una ponderación y la suma de estas ponderaciones dan un resultado final sobre 100 puntos.

2.2 Procedimiento para elaboración del tablero de control

En la Fig 2 se tiene el diagrama general del proceso de elaboración del tablero de control y a continuación, en las siguientes secciones se describe con detalle de cada una de las actividades realizadas.

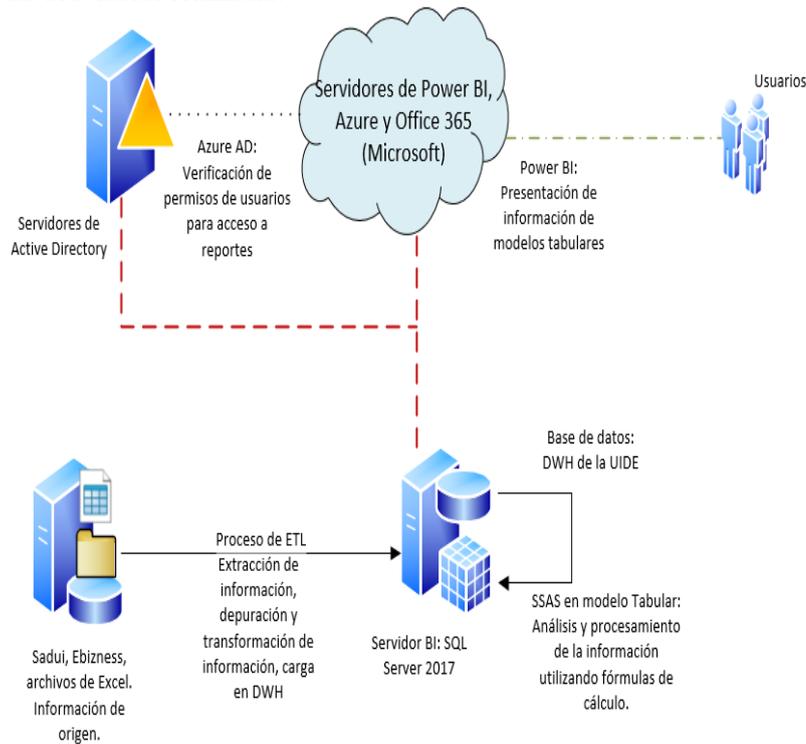


Fig 2: Diagrama general del proceso de elaboración del tablero de control

2.2.1.- Proceso de ETL para poblar el Data Warehouse (DWH)

Se creó el servidor BI en el cual se instaló el motor de datos de SQL Server que tiene el Data Warehouse, y que actualmente obtiene información de los sistemas académico (SADUI) y financiero (Ebizness), así como archivos planos e información ingresada manualmente. Los datos con los que trabaja el Data Warehouse del sistema académico son obtenidos utilizando los respaldos diarios, lo cual evita paralizar el sistema, mientras se procesa la información.

El proceso de ETL (Extracción, transformación y carga de datos) permite leer los datos de los distintos orígenes y transformarlos en información útil para el análisis (Camacho Colán, Acosta Flores, & Nuñez Huapaya, 2018). Este proceso ha permitido, entre otras cosas: Detectar información inconsistente y duplicada, estandarizar datos, eliminando algunos símbolos extraños en campos comunes, calcular el nivel de los

estudiantes para cada período lectivo, entre otras tareas. Además de la depuración de la información, el proceso de ETL permite agrupar la información del docente por años, realizando cálculos que permiten conocer su máximo nivel de título, máximo nivel de contrato, máximo nivel de titularidad, etc. Finalmente, la información es almacenada en otras tablas que son las que sirven de origen de datos para cualquier herramienta de análisis de información. (SQL Server Analysis Services, Pentaho, Power BI, Tableau, Qlik, etc).

2.2.2 Análisis y procesamiento de la información.

Para poder analizar la información se utilizó SQL Server Analysis Services (SSAS) que "es un motor de datos analíticos utilizado para el soporte a las decisiones y análisis de negocios" (Microsoft, 2018), eligiendo el uso de modelos tabulares, los cuales son compatibles tanto con el complemento Power Pivot incluido en Excel, así como con Power BI, herramienta que permite visualizar los reportes a través de la web y también a través de dispositivos móviles. Es importante indicar que para el presente trabajo se utilizó Power BI que se ubica en la sección de líderes del Cuadrante Mágico de Gartner de las plataformas para analítica e inteligencia de negocios (Microsoft, 2019).

2.2.3 Modelos Tabulares

Se creó un modelo tabular exclusivo para el análisis en base al modelo de acreditación, el cual requiere información cualitativa y cuantitativa ingresada manualmente, así como información muy particular como por ejemplo: el número de estudiantes hace dos años para análisis de retención, información del número de estudiantes en un período de origen para calcular las tasas de graduación o tasas de deserción.

2.2.4 Fórmulas

El modelo tabular utiliza un lenguaje denominado DAX, "es un lenguaje de fórmulas. Puede usar DAX para definir cálculos personalizados para las columnas calculadas y para las medidas (también conocidos como campos calculados)" (Microsoft, 2019)

El CEAACES entregó gráficas que representan las curvas de utilidad para calificar cada indicador, por ejemplo, la gráfica del indicador de posgrado en formación(Fig 3):

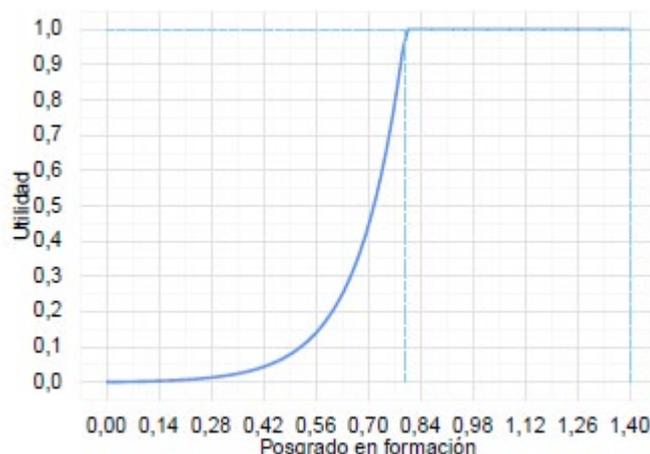


Fig 3: Curva de utilidad para el indicador Posgrado en formación

En un inicio el CEAACES no entregó la fórmula utilizada para ninguno de los indicadores, de tal manera que se realizaron fórmulas de aproximación en Excel utilizando varios puntos de referencia en el gráfico. Un poco antes de finalizar la elaboración del modelo, así como del panel de control, el CEAACES entregó un archivo que contenía las ecuaciones utilizadas para la evaluación institucional, por lo que se utilizaron estas fórmulas únicamente para los indicadores faltantes de elaborar.

Debido a que algunas fórmulas utilizadas para obtener los KPI's eran complejas, se dividieron en varias secciones parciales hasta obtener la fórmula final.

Por ejemplo, para obtener el KPI de formación posgrado, así como su aporte final al modelo de acreditación se utilizaron 5 fórmulas y 2 constantes.

Para definir una fórmula simple, se pueden realizar operaciones básicas con campos ya existentes.

Indicador Posgrado en Formación:=[Sin PHD cursando PHD o Eq.]/[Maestría o Eq.]+0.4*([Sin Maestría cursando Maestría o Eq.]/[Docentes sin PHD ni Maestría])

Indicador PF2:=IF([Indicador Posgrado en Formación]>[(Objetivo KPI) Posgrado en Formación],1,[Indicador Posgrado en Formación]/[(Objetivo KPI) Posgrado en Formación])

En el caso del KPI de posgrado en formación, antes de conocer las ecuaciones entregadas por el CEAACES, se utilizó la siguiente fórmula de aproximación:

KPI Posgrado en Formación:=IF(ISBLANK([Indicador Posgrado en Formación]),BLANK(),IF([Indicador Posgrado en Formación]>=[(Objetivo KPI) Posgrado en Formación],1,FLOOR(ABS(7.8134*POWER([Indicador PF2],6)-17.67*POWER([Indicador PF2],5)+16.82*POWER([Indicador PF2],4)-7.4574*POWER([Indicador PF2],3)+1.6*POWER([Indicador PF2],2)-0.106*[Indicador PF2]),0.0001)))

Una vez que el CEAACES entregó su ecuación de cálculo, que utilizaba exponenciales con precisión de milésimas se crearon fórmulas parciales como las siguientes:

F1RMT:=FLOOR(-0.005652+0.005652*EXP(0.232104*[Remuneración Docentes MTTP]),0.0001)
F2RMT:=FLOOR(0.9249-14.01*EXP(-0.1786875*[Remuneración Docentes MTTP]),0.0001)
F3RMT:=FLOOR(1.0267-6.418*EXP(-0.12889*[Remuneración Docentes MTTP]),0.0001)
 Finalmente, se unieron las fórmulas parciales con los límites proporcionados:
KPI Remuneración MTTP:=IF(ISBLANK([Remuneración Docentes MTTP]),BLANK(),IF([Remuneración Docentes MTTP]<=[Límite KPI 1] Remuneración MTTP],[F1RMT],IF([Remuneración Docentes MTTP]<=[Límite KPI 2] Remuneración MTTP],[F2RMT],IF([Remuneración Docentes MTTP]<=[Límite Max KPI] Remuneración MTTP],[F3RMT],1))))

2.2.5.- Power BI

Power BI es una herramienta que brinda una interfaz gráfica para la visualización de reportes de modelos tabulares, así como el análisis directo de información obtenida desde múltiples orígenes de datos.

Para la integración de los modelos tabulares con Power BI fue necesaria la instalación del componente “On-premises data gateway” en el servidor BI, el cual permite el acceso desde el sitio de Power BI. Posteriormente se crearon los reportes en la herramienta de escritorio de Power BI, y finalmente se publicaron en la web.

2.2.6.- Presentación del tablero de control

Una vez hecho el modelo tabular que permite el análisis de las distintas variables, se creó el área de trabajo Indicadores UIDE que cuenta con distintos paneles que contienen las variables necesarias para el análisis por inicio calendario, así como un panel que permite ver los resultados clasificados según las distintas categorías, y el resultado numérico de la evaluación para el año 2018.

Debido al funcionamiento de los paneles y debido a que la acreditación se debe realizar para un año específico se utilizó el año 2018.

Panel de resultados

En el panel de resultados(Fig. 4), se puede visualizar la suma del resultado de la evaluación tanto en indicadores cualitativos como cuantitativos, permitiendo ingresar tanto a la división entre la suma de indicadores cualitativos como cuantitativos, así como a los indicadores específicos en caso de no existir una subdivisión.



Fig 4: Panel de resultados- Tablero de control del modelo de acreditación

Categorías con indicadores cualitativos y cuantitativos

Al ingresar a una categoría o “criterio” que tenga componentes tanto cualitativos como cuantitativos, se mostrará un subpanel con el subtotal de cada uno de ellos. En la Fig.5 se observa los subpaneles del criterio: “Academia”.



Fig 5: Subpanel del criterio “Academia”

Paneles con KPIs Cuantitativos

Los paneles que cuentan con KPI's cuantitativos muestran un indicador tipo velocímetro junto a un tablero que muestra el color relacionado con el indicador, una línea que muestra la variación del indicador en el tiempo, y dos medidas asociadas al indicador.(Fig. 6)

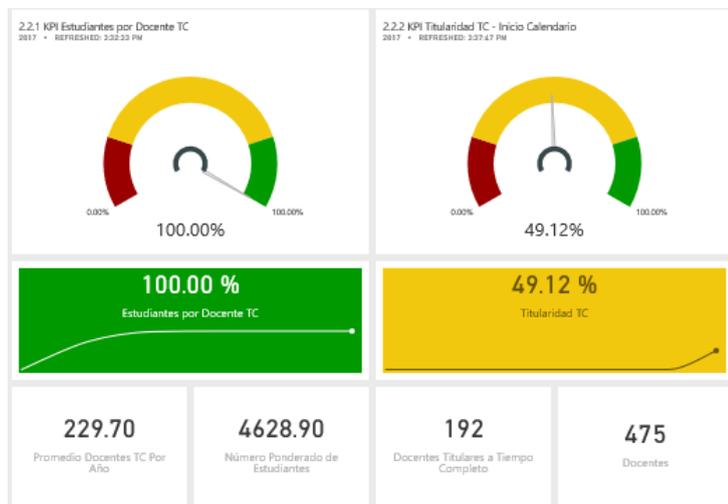


Fig 6: KPI's cuantitativos para “Estudiantes por docente TC” y “Titularidad TC”

Paneles con KPIs Cualitativos

Los paneles que cuentan con KPIs cuantitativos muestran un indicador tipo barra con la nota cualitativa asociada.(Fig 7)

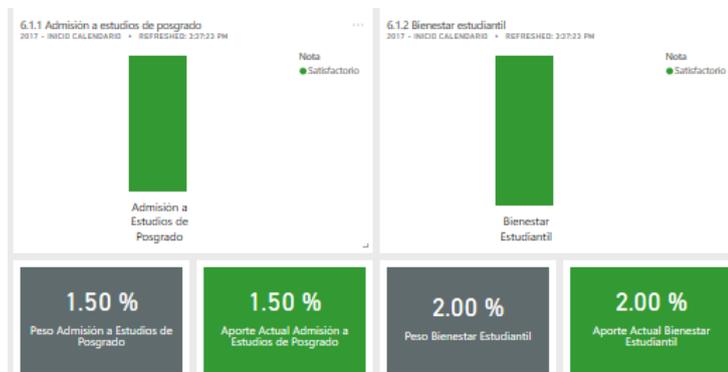


Fig 7: KPI's cualitativos para “Admisión a estudios de posgrado” y “Bienestar estudiantil”

Reportes asociados a KPIs cuantitativos

Los reportes asociados a indicadores cuantitativos, los cuales son accesibles al pulsar sobre un KPI relacionado, permiten analizar las variables utilizadas para el cálculo del KPI, así como la información histórica de dichas variables.

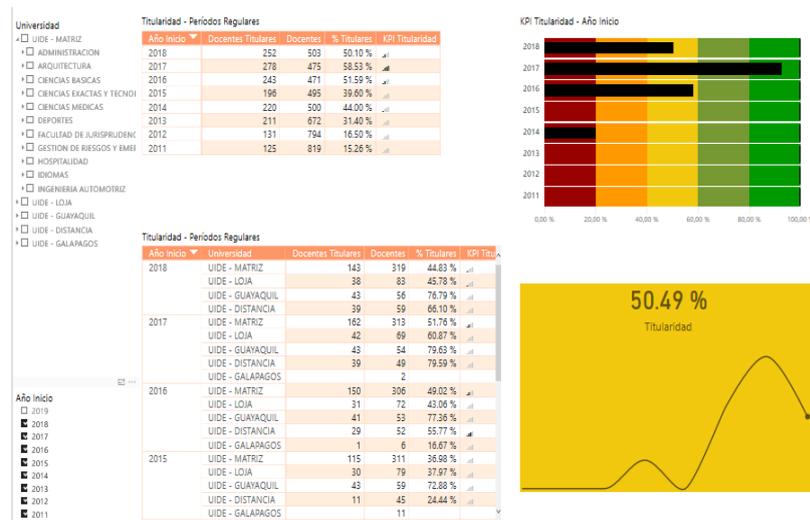


Fig 8: Reporte relacionado al KPI's “Titularidad”

2. Impacto generado

Este tablero de control está orientado a usuarios que toman las decisiones a muy alto nivel, en el caso de la UIDE es accesible únicamente por el Rector, Vicerrectores y Decanos y tuvo un alto impacto en cuanto a la facilidad que brindó para el control y seguimiento de la evolución de los indicadores del modelo de acreditación a nivel institucional. Facilitó la toma de decisiones para, por ejemplo: enfocar las nuevas inversiones, asignación de recursos, contratación de docentes, abrir nuevas carreras, cerrar carreras, cambiar la organización interna, hacer la prospectiva, entre otras.

3. Conclusiones

Elaborar un tablero de control para cualquier tipo de institución representa un reto muy importante debido a la complejidad en la selección del tipo de variables a utilizar, más aún en las universidades que son organizaciones muy particulares. Haber tomado como base el modelo de acreditación propuesto por el organismo del control del país, sin lugar

a dudas fue una gran decisión, debido a que este modelo involucra los aspectos más representativos de las universidades que deben ser considerados por las autoridades para tomar decisiones estratégicas bien sustentadas. Este proceso puede ser aplicado en cualquier universidad en cualquier país y región, ya que simplemente se debería adaptar a los modelos de acreditación vigentes en los mismos.

Un aspecto muy interesante a destacar de este tablero de control es que realiza el cálculo de indicadores en base a la información que está almacenada en los sistemas transaccionales, en este caso el sistema académico y el sistema financiero, que luego son depurados y cargados en el data warehouse a través del proceso de ETL. La información que no está disponible en estos sistemas se cargó desde archivos de Excel que tocó elaborarlos manualmente. Este fue un detonante y una excelente guía para el desarrollo de nuevos módulos, principalmente del sistema académico para que la información ya no sea ingresada manualmente.

En Ecuador en este año 2019 se lanzará un nuevo modelo de acreditación para lo cual se realizarán las modificaciones que sean necesarias al tablero control presentado en este artículo. Este trabajo sin lugar a dudas se realizará de una manera mucho más rápida debido a la base ya existente.

4. Agradecimientos

Se agradece a todas las personas que apoyaron la realización del presente trabajo en especial al Vicerrector Académico, Ec. Ramiro Canelos, por la confianza depositada y al Dr. Luis Vásquez, Decano de la Facultad de Gestión de Riesgos por haber generado la idea de la construcción de este tablero de control universitario.

5. Referencias

1. Camacho Colán, L. A., Acosta Flores, J. A., & Nuñez Huapaya, J. A. (2018). Implementación de business intelligence, utilizando la metodología de Ralph Kimball, para mejorar el proceso de toma de decisiones académicas en el Cetpro Margarita Gonzales de Danker's. Lima, Perú.
2. CEAACES. (2017). *Proceso de recategorización 2015- Informe Final*. Quito.
3. Laudon&Laudon. (2016). *Sistemas de Información Gerencial*. Mexico: Pearson.
4. Microsoft. (04 de 12 de 2018). *About SQL Server Analysis Services* . Obtenido de <https://docs.microsoft.com/en-us/sql/analysis-services/analysis-services?view=sql-server-2017>
5. Microsoft. (11 de 02 de 2019). *2019 Gartner Magic Quadrant for Analytics and Business Intelligence Platforms*. Obtenido de <https://info.microsoft.com/ww-landing-gartner-mq-bi-analytics-2019.html?LCID=EN-US>
6. Microsoft. (2019). *Expresiones de análisis de datos (DAX) en Power Pivot*. Obtenido de <https://support.office.com/es-es/article/expresiones-de-análisis-de-datos-dax-en-power-pivot-bab3fbe3-2385-485a-980b-5f64d3b0f730>

