



Tecnológico
de Antioquia
Institución Universitaria



Investigación e Innovación

en Ingeniería de Software

Volumen 3

Investigación e Innovación

en Ingeniería de Software

Volumen 3



Tecnológico
de Antioquia
Institución Universitaria

Investigación e Innovación en Ingeniería de Software / volumen 3 Fabio Alberto Vargas Agudelo, Darío Enrique Soto Duran, Juan Camilo Giraldo Mejía, compiladores. – Medellín : Publicar T, Sello Editorial TdeA, 2019.

Corrección de estilo, diseño, diagramación y animación

Divegráficas S.A.S.
divegraficas@gmail.com

Derechos reservados del Tecnológico de Antioquia –
Institución Universitaria

“Los artículos publicados en este libro incorporan contenidos derivados de procesos de investigación y reflexión académicas no representan necesariamente los criterios institucionales del Tecnológico de Antioquia. Los contenidos son responsabilidad exclusiva de sus autores. Obra protegida por el derecho de autor. Queda estrictamente prohibida su reproducción, comunicación, divulgación, copia, distribución, comercialización, transformación, puesta a disposición o transferencia en cualquier forma y por cualquier medio, sin la autorización previa, expresa y por escrito de su titular. El incumplimiento de la mencionada restricción podrá dar lugar a las acciones civiles y penales correspondientes. © 2019 Todos los derechos de autor reservados.”

Hecho en Colombia





Investigación e Innovación en Ingeniería de Software
Tecnológico de Antioquia – Institución Universitaria

Publicar T – Sello Editorial TdeA
Tecnológico de Antioquia – Institución Universitaria
Calle 78B No. 72A – 220
PBX: (+57)(4) 444 37 00
www.tdea.edu.co
Medellín - Colombia

Dr. Lorenzo Portocarrero Sierra
Rector

Dr. Fabio Alberto Vargas Agudelo
Director de Investigación – Coordinación del Sello Editorial

Dr. Darío Enrique Soto Durán
Decano de la Facultad de Ingeniería

Dr. Lorenzo Portocarrero Sierra

Rector

Dr. Fabio Alberto Vargas Agudelo

Director de Investigación – Coordinación del Sello
Editorial

Dr. Dario Enrique Soto Duran

Decano de la Facultad de Ingeniería

Dr. Juan Camilo Giraldo Mejía

Lider de Grupo GIISTA

Autores:

Carmen Janeth Parada

María del Pilar Rojas Puentes

Judith del Pilar Rodríguez

Fabio Alberto Vargas Agudelo

Juan Camilo Giraldo Mejía

Dario Enrique Soto Durán

Jaime Alberto León Rincón

Matías Herrera Cáceres

Milton Jesús Vera Contreras

Óscar Alberto Gallardo Pérez

Alexander Barinas López

John Alexander Bohada Jaime

Iván Andrés Delgado González

Fredy Alonso Vidal Alegría

Víctor Daniel Gil Vera

Catalina Quintero López

Diana Maria Montoya Quintero

Sonia Jaquelliny Moreno Jiménez

Mónica María Rojas Rincón

Alejandro Peláez Piedrahita

Paula Andrea Vargas Álvarez

Adriana Xiomara Reyes Gamboa

Sara Cristina Vera Sepúlveda

Mauricio Escobar Tobón

Sandra Mateus Santiago

Juan Carlos Valdés Quintero

Alexandra Urán

John Fernando Escobar Martínez

Vanessa Paredes Zuñiga

Esteban Zapata Trejos

Juan Camilo Parra Toro

Wilson Alberto Contreras Espinoza

José Orlando Maldonado Bautista

Karime Alejandra Gómez Bastidas.

Maria Camila Dueñas Cuellar

ISBN: 978-958-52397-5-3

Investigación e Innovación en Ingeniería de Software
/ volumen 3 Fabio Alberto Vargas Agudelo, Darío
Enrique Soto, Juan Camilo Giraldo, compiladores. –
Medellín : Publicar T, Sello Editorial TdeA, 2019.

7	PREFACIO
9	Capítulo I ANÁLISIS DE LAS METODOLOGÍAS ÁGILES PARA EL DESARROLLO DE SOFTWARE
17	Capítulo II GUÍA TÉCNICA PARA LA ESPECIFICACIÓN DE REQUISITOS DE SOFTWARE EN EL DESARROLLO DE APLICACIONES PARA CIUDADES INTELIGENTES
28	Capítulo III PRUEBAS AUTOMÁTICAS PARA EVALUAR CURSOS DE PROGRAMACIÓN DE COMPUTADORES
43	Capítulo IV CRITERIOS Y MÉTRICAS PARA EVALUAR LA SEGURIDAD EN APLICACIONES WEB: METODOLOGÍA MESW
54	Capítulo V CARACTERIZACIÓN Y DISEÑO DE UN MODELO DE DATOS PARA LA GESTIÓN DE INDICADORES DEL SISTEMA DE ASEGURAMIENTO DE ALTA CALIDAD DE LA INSTITUCIÓN UNIVERSITARIA COLEGIO MAYOR DEL CAUCA
62	Capítulo VI AGRUPAMIENTO JERÁRQUICO SOBRE USO DE LAS TIC EN COLOMBIA: UN CASO DE APLICACIÓN DEL MACHINE LEARNING
73	Capítulo VII OBTENCIÓN DEL CONOCIMIENTO PARA ANÁLISIS DE REQUISITOS EN EL DESARROLLO DEL SOFTWARE
84	Capítulo VIII ANALÍTICA DE DATOS PARA DETERMINAR LA AFINIDAD ACADÉMICA Y PERSONAL EN PROGRAMAS DE INGENIERÍA
96	Capítulo IX FACTORES QUE INCIDEN EN EL FUNCIONAMIENTO DE UNA PMO EN LAS EMPRESAS DE TI: UNA REVISIÓN DE LITERATURA

103	Capítulo X ESTIMACIÓN DE PRECIPITACIÓN Y EVAPOTRANSPIRACIÓN UTILIZANDO SISTEMAS DE INFORMACIÓN GEOGRÁFICA A PARTIR DE DATOS OBTENIDOS EN ESTACIONES METEOROLÓGICAS UBICADAS EN LA CUENCA DEL RÍO NEGRO
113	Capítulo XI PLATAFORMAS DE SENSORES REMOTOS EN LA GENERACIÓN DE CARTOGRAFÍA COMO INSTRUMENTO DE APOYO A LA GESTIÓN DE ECOSISTEMAS ESTRATÉGICOS
120	Capítulo XII MODELAMIENTO DE LA CONCENTRACIÓN DE CONTAMINANTES EN SUELOS POR EL MÉTODO DE DIFERENCIAS FINITAS UTILIZANDO ESQUEMAS DE PRIMER Y SEGUNDO ORDEN

PREFACIO

La tercera edición del libro *Investigación e innovación de la ingeniería de software* compila los resultados de investigación que relacionan las líneas: Seguridad de la Información, Ingeniería de Software, y Gestión del Conocimiento. Son doce capítulos los que conforman esta tercera edición, con un formato que facilita la lectura, cuyo propósito es enriquecer a los lectores con la diversidad de contribuciones investigativas en diferentes tópicos. El libro presenta de forma estructurada los resultados de cada una de las investigaciones articuladas con la industria 4.0, y recoge, más que una tendencia, los temas de interés y aplicación tanto en el contexto académico como organizacional. Se resaltan los procesos soportados por la analítica de datos, la estadística, las bases de datos, la gestión de la información y del conocimiento, la ingeniería de software, al igual que la seguridad de la información. Son todas estas tecnologías, herramientas y metodologías las que permiten el mejoramiento continuo de los procesos, a través del monitoreo y la integración con otros sistemas específicos de las organizaciones que posibilitan la transferencia y la apropiación del conocimiento en contextos de investigación tanto a nivel de pregrado como de posgrado.

Estamos seguros de que los resultados en este nuevo volumen se convierten en insumo para la formulación y el desarrollo de nuevos proyectos de investigación, de los cuales se generarán nuevas contribuciones que podrán ser los capítulos para la cuarta edición del libro de ingeniería de software.

Fabio Alberto Vargas Agudelo
Juan Camilo Giraldo Mejía
Darío Enrique Soto Durán
Compiladores

Capítulo I

Análisis de las Metodologías Ágiles para el Desarrollo de Software

Carmen Janeth Parada - janethpc@ufps.edu.co

Docente investigadora de GIDIS, Universidad Francisco de Paula Santander, Cúcuta

Pilar Rojas Puentes - pilarrojas@ufps.edu.co

Docente investigador de GIDIS, Universidad Francisco de Paula Santander, Cúcuta

Judith del Pilar Rodríguez - judithdelpilarrrt@ufps.edu.co

Docente investigadora de GIDIS, Universidad Francisco de Paula Santander, Cúcuta

I. INTRODUCCIÓN

Las metodologías ágiles han tomado gran importancia e interés en los equipos de trabajo de la industria del software, por las ventajas y bondades que se obtienen con su uso en el desarrollo rápido, la flexibilidad a los cambios a que están expuestos este tipo de proyectos, la entrega de productos a satisfacción del cliente y la adaptabilidad de estas metodologías.

El artículo describe las características que tienen las metodologías ágiles, para facilitar la estructuración de una nueva metodología que permita a los equipos de trabajo personalizarlas, de acuerdo con los proyectos que desarrollan. Se responde a la pregunta: ¿Cómo especificar una nueva metodología ágil a partir de las características de las metodologías ágiles existentes?

La estructura del presente artículo se organizó de la siguiente manera: en el apartado 2 se presentan los conceptos más relevantes en cuanto a metodologías ágiles en los beneficios, y el contexto de aplicación e impacto; en el apartado 3, la metodología de desarrollo de la investigación; en el apartado 4, el análisis de las características de las metodologías ágiles; en el apartado final, la conclusión y las referencias usadas en la investigación.

II. REVISIÓN DE LITERATURA

A. Metodologías Ágiles para el Desarrollo de Software

El Instituto Nacional de Tecnologías de la Comunicación de España (Inteco) [1], en su capítulo de Ingeniería de Software, define “una metodología como un conjunto integrado de técnicas y métodos que permite abordar de forma homogénea y abierta cada una de las actividades del ciclo de vida de un proyecto de desarrollo”. Igualmente, en [2] se define que, en el ámbito de los proyectos, “una metodología se considera un proceso ordenado compuesto de pasos e involucra herramientas, técnicas, responsabilidades y entregables, el cual pasa por un estado de avance, concluye y se entrega con éxito”. Lo anterior permite indicar que una metodología establece un camino para desarrollar un proyecto de manera sistemática, y proporciona un estándar de trabajo al equipo de desarrollo para alcanzar los objetivos propuestos y satisfacer los requisitos de los interesados.

Las metodologías de desarrollo presentan dos enfoques: predictivos y adaptativos [3]; los enfoques predictivos son utilizados cuando se tienen identificados los requisitos con anterioridad, antes de iniciar el proceso de desarrollo y cuando el producto software final no está sujeto a cambios. En cambio, en los enfoques adaptativos, el alcance del producto se define y se aprueba antes del comienzo de una iteración.

Bajo esta definición nacieron las metodologías ágiles, las cuales son adaptativas más que predictivas, y se centran más en las personas que en los procesos [4]. Con base en estas características, a inicios del año 2001, un grupo de desarrolladores y de expertos de la industria de software consolidaron el término ágil aplicado al desarrollo de software y definieron el “Manifiesto Ágil” [5], donde plasmaron los principios y valores que los grupos desarrolladores deben aplicar para crear softwares rápidamente y respondiendo a los cambios que puedan surgir a lo largo de un proyecto. Estas metodologías se caracterizan principalmente por representar la antítesis de lo que es el tradicional proceso para desarrollar softwares; pasan por alto la utilización de elaborados casos de uso, la exhaustiva definición de requerimientos y la producción de una extensa documentación [6].

Para los proyectos pequeños y medianos, donde los requerimientos cambian constantemente y donde no existe la necesidad de tener equipos grandes, los usos de metodologías ágiles permiten obtener productos softwares en tiempos más cortos y con costos reducidos sin perder la calidad del producto desarrollado [7].

B. Beneficios del Uso de las Metodologías ágiles

Usar metodologías ágiles para el desarrollo de softwares no solo permite disminuir los tiempos de desarrollo, sino que se pueden dar otros beneficios como simplificación de la sobrecarga de procesos, calidad mejorada, mejora en la previsibilidad con una adecuada gestión del riesgo, mejor perfil de productividad, entre otros.

Simplificación de la sobrecarga de procesos. Es normal que los equipos de trabajo que desarrollan softwares usando metodologías tradicionales tengan sobrecarga de trabajo para poder cumplir con el desarrollo del software, de acuerdo con los estándares de la industria. Las metodologías ágiles permiten cumplir los estándares definidos por la industria con menos sobrecarga utilizando iteraciones más cortas y empaquetadas, lo que genera

beneficios en un proceso que puede adaptarse a los cambios que surgen, y donde se requiere menos sobrecarga en el proceso desde su inicio hasta su finalización, lo que implica menos trabajo a medida que se acerca la fecha final.

Calidad mejorada. Las prácticas de desarrollo ágil proporcionan la mínima funcionalidad con la máxima calidad. La mínima funcionalidad no implica necesariamente una pobre funcionalidad, sino la suficiente como para conseguir que el trabajo se realice. Desarrollar iteraciones en poco tiempo y demostrar a los clientes los productos pronto y con frecuencia, permite tanto a los clientes como a los equipos de desarrollo ponerse de acuerdo en que el producto cumple con cada una de sus necesidades.

Mejorar la previsibilidad a través de una adecuada gestión del riesgo. Hay muchas razones justificadas por las cuales los equipos desarrolladores no cumplen con las entregas de un producto en las fechas determinadas; sea cual sea la causa que genera este incumplimiento, los negocios requieren que estas fechas de entrega se cumplan. Las metodologías ágiles ayudan a que los proyectos de tecnología cumplan con las fechas de entrega estipuladas en los contratos; esto se puede hacer de dos formas: 1) Priorizando los riesgos que se puedan causar en el desarrollo de un software y, 2) Evaluando el riesgo en paralelo, esto es, varios equipos trabajan de forma paralela resolviendo el mismo problema con diferentes soluciones, lo cual puede permitir tomar decisiones claves para seguir el desarrollo.

Mejor perfil de productividad. Los equipos ágiles tienden a ser muy productivos desde la primera iteración hasta el lanzamiento y ritmo, se tienen que gestionar de modo que no se produzca agotamiento. Los equipos ágiles que mantienen este código de trabajo con cada iteración, permiten realizar pruebas de rendimiento y sistemas desde el principio, es decir, en las primeras iteraciones. De este modo, defectos críticos, como problemas de integración, se descubren antes, la calidad general del producto es mayor y el equipo funciona de manera más productiva durante todo el ciclo de desarrollo.

C. Aplicación del Agilismo en un Contexto Académico

El estudio sobre las experiencias de formación en metodologías ágiles, realizado en los años 2015-2016 con los estudiantes del curso de Análisis y Diseño de Sistemas e Ingeniería del Software del Programa de Ingeniería de Sistemas (UFPS) [8], se apoyó en una investigación de campo; uno de los aspectos tenidos en cuenta tuvo que ver con la validez para garantizar la calidad del estudio. En la ejecución del procedimiento metodológico cuantitativo se aplicó la prueba piloto, cuestionario de satisfacción ya validada a estudiantes que no formaban parte de la muestra, pero que presentaban las mismas características de los sujetos evaluados. Para hallar el coeficiente de confiabilidad se procedió de la siguiente manera:

- Aplicación de la prueba piloto a un grupo de 30 sujetos (estudiantes) pertenecientes a la muestra de estudio, con características equivalentes a la misma.
- Codificación de las respuestas; transcripción de las respuestas en una matriz de tabulación de doble entrada con el apoyo del equipo estadístico de la Universidad.
- Interpretación de los valores tomando en cuenta la escala de Likertt.
- Determinación de resultados con tabulación simple y tabulación cruzada.

Conocido el método de investigación, se procedió a desarrollar el estudio, y al iniciar cada semestre académico, en el aula de clase se impartieron los conceptos y talleres para resolver un problema real, y se organizaron los trabajos por equipos conformados por cuatro o cinco estudiantes.

Durante el desarrollo del proyecto se organizaron sesiones semanales para el trabajo en equipo, la documentación y el desarrollo. En estos proyectos se evidenció que los estudiantes ven de manera positiva y favorable no sólo la estrategia implementada, sino el hecho de poder aplicar una metodología de desarrollo ágil en un caso práctico. Para los estudiantes fue un aspecto positivo, y se destaca que

el trabajo realizado les permitió aplicar las mejores prácticas del desarrollo de software, así como hacer el seguimiento al proyecto (donde el cliente y el profesor guían el trabajo).

Lo anterior permitió conocer no sólo apreciaciones de los estudiantes, sino la aplicación de las mejores prácticas de la ingeniería de software y el uso de metodologías ágiles en proyectos reales; era importante resolver el problema que se venía presentando en cuanto a unificar los conceptos teóricos con la práctica, aspecto que se había evidenciado en semestres anteriores.

Las metodologías XP y SCRUM fueron las más seleccionadas (75%); para los estudiantes era atractivo mejorar y agilizar el proceso de desarrollo del software. Sin embargo, consideraban que se presentaban conflictos entre los miembros del equipo, bien por el ritmo de trabajo o por capacidades, al responder con el trabajo efectuado. Igualmente, la exigencia era mayor si no conocían de antemano el proceso de desarrollo. Asimismo, durante el proceso de desarrollo los estudiantes presentaron dificultades en cuanto a las pruebas, el *refactoring* e incluso al integrar las pruebas, donde la experiencia y la disciplina fueron aspectos relevantes.

Otro aspecto positivo del estudio fue el uso de las metodologías ágiles. En el curso de Ingeniería de Software se evidenció que, en el mantenimiento correctivo de los proyectos realizados en el contexto planteado, no se presentaban los requerimientos esperados por los clientes, y además los hitos se retrasaban, lo que ocasionaba mayor presión al equipo de desarrollo, al exigírsele mejorar los tiempos y mantener los hitos. Por otro lado, se observó que el cliente no estaba satisfecho con la entrega de las mejoras y/o correcciones, ya que se entregaban en fechas posteriores a las pactadas.

Para solucionar y mejorar los aspectos anteriormente descritos, se determinó implementar la metodología SCRUM en los proyectos de curso de las asignaturas Análisis y Diseño e Ingeniería de Software, esto permitió mejorar la satisfacción del cliente al informarle sobre los avances realizados

y evitar así malentendidos; mantener motivados a los equipos de desarrollo por los resultados obtenidos; conocer los avances del proyecto cuando los interesados lo solicitaran; ahorrar tiempo y costos cuando se presentaran cambios en los requisitos del producto, y cumplir con los tiempos establecidos para la entrega.

III. METODOLOGÍA

Para responder a las preguntas de investigación, se condujo un estudio descriptivo exploratorio, el cual permitió conocer las percepciones y las prácticas de los profesionales que desarrollan software, no se limitó sólo a recolectar los datos, sino a identificar las relaciones entre dos o más variables. Posteriormente, se recogieron los datos sobre la base de la hipótesis y se resumió de manera cuidadosa la información para ser analizada con el fin de entregar resultados que contribuyeran al estudio.

Luego se aplicó el instrumento de la encuesta, que fue diligenciada por los desarrolladores pertenecientes a la población objetivo. Las respuestas complementaron los hallazgos del estudio. La encuesta contenía preguntas relacionadas con las prácticas ágiles de las diferentes metodologías existentes, su utilización, sus características y sus ventajas en el desarrollo del software. Luego se procedió a realizar el análisis de los datos, con la técnica de comparación de conceptos, y se intentó crear una metodología alrededor de las lecciones aprendidas por los investigadores y de los resultados del proceso de desarrollo con el fin de adoptar una metodología ágil, y de esta manera responder a las preguntas de investigación formuladas. Por último, se consolidaron las conclusiones encontradas en el estudio.

Recolección de datos: En el artículo se señalan los datos obtenidos y la naturaleza de la población de la cual fueron extraídos. La población fue de 1.234 sujetos con una muestra de 104 sujetos (profesionales egresados del programa de Ingeniería de Sistemas de la UFPS) [9]. Una vez identificada la población, se decidió si se recogían los datos de la población total o de una muestra representativa de

ella. El método se eligió en función de la naturaleza del problema y de la finalidad para utilizar los datos.

Investigación descriptiva: El instrumento tipo encuesta se diseñó en dos secciones, y permitió encontrar soluciones que surgieron al aplicarlo. Se efectuó una descripción minuciosa de la temática por tratar, a fin de justificar las percepciones de los desarrolladores de software y las prácticas que efectúan. Su objetivo no se limitó solo a determinar el problema, sino también a comparar entre las diferentes metodologías ágiles, su uso y aplicación en el proceso de desarrollo. El estudio permitió determinar un alcance de acuerdo con la población objetivo y la información recogida, se realizó el número de secciones del instrumento en relación con el objeto de estudio y el problema dado.

La muestra calculada es probabilística simple, en ella todos los individuos tienen la misma probabilidad de ser elegidos para conformarla y, consiguientemente, las posibles muestras de tamaño no tienen la misma probabilidad de ser seleccionadas. Este método de muestreo probabilístico nos asegura la representatividad de la muestra extraída [10].

IV. RESULTADOS

Una de las herramientas más importantes y puesta en práctica en la mayoría de las organizaciones exitosas del mundo de hoy es la incorporación de metodologías para la gestión de los proyectos. En [2] se indica que la aparición de metodologías ágiles ha sido constante desde cuando se planteó el concepto. Las metodologías ágiles se emplean para la gestión de proyectos de cualquier disciplina, y han sido apropiadas por la industria del software. Se mencionan las metodologías ágiles más utilizadas por los equipos de trabajo en el desarrollo de software [10].

Como se observa en la Tabla 1, las metodologías ágiles más utilizadas por los equipos de trabajo son Scrum (43,42 %), XP (10,50 %) y Crystal (9,20 %), las cuales se toman en cuenta en los resultados del presente artículo.

Tabla 1. Metodologías ágiles

Metodología	Porcentaje resultado del estudio
Adaptive Software Development (ASD)	5,3
Scrum	43,4
Cristal Methods (CM)	9,2
DSMD	3,9
eXtreme Programming (XP)	10,5
Feature-Driven Development (FDD)	1,3
Lean Development (LD)	3,9
Otras	2,6

Fuente: Elaborado a partir de [9], [10]

Las metodologías ágiles son un marco de trabajo de la ingeniería de software que define iteraciones en el desarrollo a lo largo del ciclo de vida del proyecto. La mayoría de las metodologías ágiles pretenden minimizar los riesgos de desarrollo del producto en corto tiempo. En la Tabla 2 se especifican las variables analizadas en el estudio realizado en [10].

Tabla 2. Categoría y variables identificadas en las metodologías ágiles

Categoría	Variable (características)
Generales	Ciclo de vida
	Entregas parciales
Equipo de trabajo	Estilo de comunicación
	Tamaño
	Jerarquía del equipo
Gestión del proyecto	Cultura de gestión
	Planificación
	Retroalimentación
El cliente	Disponibilidad del cliente
	Acceso a usuarios expertos
Procesos y herramientas	Tiempo de la iteración
	Adaptabilidad

Fuente: Tomado de [10].

En la Tabla 3 se describen las características evaluadas en las metodologías Scrum, XP y Cristal.

Tabla 3. Evaluación de características en XP, Scrum y Crystal

Características	XP	SCRUM	CRISTAL
Ciclo de vida	Iterativo e incremental	La metodología seleccionada para el desarrollo del producto es la que determina el flujo de trabajo en el <i>sprint</i>	Iterativo e incremental
Entregas parciales	Sujetas a cambios durante el proyecto	No están sujetas a cambios	Realiza entregas frecuentes:
Comunicación equipo	Comunicación fluida entre los miembros del equipo	Comunicación fluida entre los miembros del equipo	Osmótica
Tamaño equipo	Entre 3 - 10	Max 8	Depende de la complejidad y tamaño del proyecto.
Jerarquía del equipo	No existe	Pequeña jerarquía definida.	Estructurada
Cultura de gestión	No existe	Existe	Existe
Planeación	No	Sí	Sí
Retroalimentación	Sí	Sí	Sí
Disponibilidad del cliente	Permanente	Permanente	Periódico

Características	XP	SCRUM	CRISTAL
Acceso a usuarios expertos	Comunicación continua	Integración permanente del experto	Reuniones con periodicidad mayor
Adaptabilidad	Previsión en el diseño del software	A los cambios entre iteraciones	A los cambios entre iteraciones
Tiempo de la Iteración	De 1 a 3 semanas	De 1 a 4 semanas, conocidas como <i>Sprint</i>	Depende cómo se definan las entregas

Fuente: Elaboración propia.

A continuación, se realiza el análisis de las características según las categorías definidas en la Tabla 2.

Generales. Las tres metodologías tienen un ciclo de vida interactivo e incremental, y se diferencian particularmente en el número de etapas que cada una de ellas tiene definidas:

a) XP: Desarrolla el producto basado en seis (6) etapas: Exploración, Planificación de la Entrega, Iteraciones, Producción, Mantenimiento y Muerte del Proyecto. Esta metodología se caracteriza por hacer entregas parciales al finalizar la iteración, la cual puede estar sujeta a cambios en las siguientes iteraciones.

b) Scrum: Su ciclo de vida contempla cuatro (4) etapas: Planificación del *sprint*, Etapa de desarrollo, Revisión del *sprint* y Retroalimentación. A diferencia de XP, Scrum se caracteriza por realizar entregables en los *sprints*, los cuales son aprobados por el cliente y no están sujetos a cambios en los siguientes *sprints*.

c) Crystal: El número de etapas del ciclo de vida depende del proyecto; emplea una planeación adaptativa que se fundamenta en el proceso de construir el producto teniendo en cuenta dos factores importantes: la complejidad y el tamaño del producto. Se caracteriza por que al finalizar cada iteración entrega un producto que puede ser funcional para el cliente o un insumo para la siguiente iteración.

Equipo de trabajo. Para las tres metodologías los equipos de trabajo son importantes. El estilo de comunicación de XP, Scrum y Crystal es fluido, mantiene una comunicación constante, rápida y oportuna entre los miembros del equipo de trabajo. A diferencia de XP y Scrum, Crystal tiene un alto costo de *feedback*, debido a la reiteración de la ejecución de los procesos de acuerdo con la evaluación de los entregables. El tamaño de los equipos de trabajo en XP oscila entre 3 a 10 miembros, en Scrum entre 3 y 8, y en Crystal [11] depende de la complejidad y el tamaño del proyecto: de 1 a 6, 7 a 20, 21 a 40, 41 a 80 y de 81 a 200 personas. En cuanto a la jerarquía del equipo de trabajo, en XP no se define una estructura porque todos los miembros del equipo están en un solo nivel para ejercer la toma de decisiones, lo contrario de Scrum y Crystal; Scrum (referencia) tiene una jerarquía definida en: *stakeholders*, *Product owner*, Scrum máster, equipo autoorganizado, y en Crystal la estructura del equipo depende de la cantidad de miembros que lo componen.

Gestión del proyecto. De las tres metodologías analizadas, en XP no se evidencia cultura de gestión, por el contrario, Scrum y Crystal presentan un grupo de procesos como planificación, ejecución, seguimiento y control que permiten la administración del proyecto, y Crystal tiene además un proceso para hacer cierre. Aunque XP no tiene cultura de gestión, algunos equipos de trabajo planean actividades superficialmente, lo contrario ocurre con Scrum y Crystal, que desarrollan una planeación exhaustiva. Las tres metodologías hacen retroalimentación, la cual es costosa en Crystal por el hecho de hacer el proceso reiteradamente.

El cliente. Para XP y Scrum, la disponibilidad del cliente en el desarrollo del producto debe ser permanente, mientras que para Crystal el contacto con el cliente se puede hacer periódicamente dependiendo de la magnitud del proyecto en tiempo y complejidad. Igualmente, en XP y Scrum la comunicación del equipo de trabajo con los expertos requeridos para el desarrollo del producto debe ser continua, mientras que en Crystal no se requiere al experto de forma permanente sino en periodos definidos.

Procesos y herramientas. El tiempo de la iteración en XP oscila entre 1 a 3 semanas; Scrum determina que las iteraciones son de máximo 4 semanas; mientras que en Crystal el tiempo de la iteración depende de cómo se hayan definido las entregas parciales de los entregables. Respecto a la adaptabilidad de los procesos para ejecutar los cambios, XP permite cambios a partir del diseño; Scrum y Crystal permiten cambios entre las iteraciones.

V. CONCLUSIONES

En un proyecto de desarrollo de software no existe una metodología universal, la elección de ésta depende de la complejidad y el tamaño del proyecto. La elección de metodologías ágiles como XP y Scrum no depende de la complejidad y el tamaño del proyecto, mientras que para Crystal estas dos características son fundamentales.

Al hacer las comparaciones de las metodologías descritas en el artículo en cuanto a las etapas que cubren el ciclo de vida, las tres metodologías seleccionadas basan su desarrollo en un modelo iterativo/incremental, siendo Scrum la que utiliza el concepto de *sprint*.

Las iteraciones definidas en las tres metodologías tienden a desarrollarse en periodos fijos de tiempo; los cambios constantes son permitidos sólo en XP. En ellas, el cliente y el experto tienen constante comunicación con el equipo de trabajo, excepto en Crystal que mantiene una periodicidad mayor en la comunicación.

El tamaño del equipo de trabajo tiene una relación directa con el tamaño del proyecto.

Finalmente, los resultados obtenidos del análisis de las características de las metodologías ágiles, permiten a los líderes de los equipos de desarrollo estructurar una metodología que se ajuste a las necesidades del proyecto, teniendo en cuenta los activos del negocio y los factores ambientales.

REFERENCIAS

- [1] Instituto Nacional de Tecnologías de la Comunicación (Inteco), *Ingeniería del Software: metodologías y ciclos de vida*. España: Laboratorio Nacional de Calidad de Software, 2009. Disponible en: <http://www.inteco.es/> (Accedido el 10 de enero de 2014).
- [2] H. F. Castro, T. Velásquez and M. P. Rojas. "Adoption of project management methodologies in Colombia project manager's perspective", *J. Phys. Conf. Ser.*, vol. 1126, Conference 1, 2018. Available: <https://iopscience.iop.org/article/10.1088/1742-6596/1126/1/012032>
- [3] Project Management Institute (PMI), *A Guide to the Project Management Body of Knowledge (PMBOK® Guide) – Sixth Edition, Official Spanish Translation*, Paperback PMI, 2017.
- [4] M. Fowler, *The New Methodology*, Technical report, 2005. [Online]. Available: <http://www.martinfowler.com/articles/newMethodology.html>
- [5] L. E. Gimson Saravia, *Metodologías ágiles y desarrollo basado en conocimiento*. Buenos Aires: Universidad Nacional de la Plata, Facultad de Informática, 2012.
- [6] A. English, "Extreme Programming: It's Worth a Look", *IEEE IT Pro*, vol. 4, No. 3, pp. 48-50, May/June 2002.
- [7] S. Ambler, *Agile Modeling: Effective Practices for eXtreme Programming and the Unified Process*. New York: John Wiley & Sons, Inc., 2002.

- [8] J. P. Rodríguez, “Estudio sobre las experiencias de formación en metodologías ágiles realizado en los años 2015-2016 en los cursos de Análisis y Diseño de Sistemas e Ingeniería del Software del programa de Ingeniería de Sistemas de la Universidad Francisco de Paula Santander”. Documento de Trabajo, 2017.
- [9] C. J. Parada, M. P. Rojas and F. H. Vera, “Study of the use of agile methodologies in the development of software construction projects in Colombia”, *J. Phys. Conf. Ser.*, vol. 1126, Conference 1, 2018. Available: <https://iopscience.iop.org/article/10.1088/1742-6596/1126/1/012056/pdf>
- [10] C. J. Parada, “Primer Informe Técnico Proyecto Metodología Ágil aplicada a los proyectos de Ingeniería de Sistemas de la UFPS que involucre el desarrollo de Aplicaciones Móviles”, 2014.
- [11] A. Cockburn, *Agile software development*. Addison Wesley, 2001.

Capítulo II

Guía Técnica para la Especificación de Requisitos de Software en el Desarrollo de Aplicaciones para Ciudades Inteligentes

Fabio Alberto Vargas Agudelo - fvargas@tdea.edu.co

Docente-investigador, Tecnológico de Antioquia - I. U., Medellín, Colombia

Jaime Alberto León Rincón - jimmyleon7@gmail.com

Docente-investigador, Tecnológico de Antioquia - I. U., Medellín, Colombia

Darío Enrique Soto Durán - dsoto@tdea.edu.co

Docente-investigador, Tecnológico de Antioquia - I. U., Medellín, Colombia

Juan Camilo Giraldo Mejía - jgiraldo1@tdea.edu.co

Docente-investigador, Tecnológico de Antioquia - I. U., Medellín, Colombia

I. INTRODUCCIÓN

Las ciudades inteligentes (*Smart cities*) son un fenómeno que crece con el avance de la tecnología, usan componentes tecnológicos y se constituyen de diversas características que las definen. Para que dichas ciudades sean una realidad se requiere un conjunto de elementos entre los que se destacan el hardware y el software, siendo el software uno de los componentes principales y complejo a la hora de desarrollarlo, si se tiene en cuenta que para su funcionamiento tiene que interactuar con diversos componentes e infraestructuras tecnológicas. Esto plantea grandes retos para los desarrolladores en la especificación de los requisitos, lo que obliga a explorar y proponer nuevas formas de llevar a cabo esta actividad.

Se evidencia un conjunto limitado de metodologías para la fase de especificación de requisitos en el desarrollo de estas aplicaciones,

por tal motivo se hace necesario proponer una guía técnica para esta fase, la cual influirá de manera positiva en las fases posteriores del desarrollo de software, para lograr un producto de calidad que satisfaga las necesidades cambiantes y exigentes de los habitantes de estas ciudades.

En el capítulo se presenta una guía técnica para la especificación de requisitos de software para ciudades inteligentes, que recoge los elementos y características esenciales de este tipo de aplicaciones, procurando una especificación de requisitos pertinente y no ambigua. Todo ello basado en un conjunto de formatos que compendian la información necesaria en este tipo de aplicaciones.

El capítulo se presenta de la siguiente manera: en la sección 1 se contextualiza el término *ciudad inteligente*; en la sección 2, se busca generar un inventario con los componentes y características necesarios para el desarrollo de aplicaciones en

ciudades inteligentes; en la sección 3, se presentan metodologías de ingeniería de requisitos utilizadas en aplicaciones para ciudades inteligentes; en la sección 4 se exponen los resultados, mediante la guía técnica para la especificación de requisitos funcionales y un ejemplo de aplicación; por último, se presentan las conclusiones.

Este capítulo es producto de la investigación para trabajo de grado del estudiante Jaime Alberto León Rincón, para optar al título de Ingeniero en Software, y de la experiencia investigativa de los autores en el área.

II. METODOLOGÍA

Este trabajo se llevó a cabo a partir de una revisión de literatura sobre el concepto de ciudad inteligente, el inventario de componentes y características propias de este tipo de aplicaciones, y el análisis de metodologías de ingeniería de requisitos aplicables para ciudades inteligentes.

A. Ciudad Inteligente

Según [1], las ciudades inteligentes responden a proyectos piloto basados en el uso de las TIC, para la transformación de la ciudad en un entorno más habitable y competitivo a nivel mundial.

Según [2], las ciudades inteligentes son aquellas que aplican las Tecnologías de la Información y la Comunicación (TIC), en procura de que sus habitantes mejoren la calidad de vida, y alcancen

un desarrollo sostenible utilizando sus recursos de la mejor forma posible.

De acuerdo con [3], una ciudad es inteligente cuando al menos dispone de una iniciativa que aborde las siguientes dimensiones: *Smart Economy*, *Smart People*, *Smart Mobility*, *Smart Environment*, *Smart Governance* y *Smart Living*, mediante el uso de las tecnologías para mejorar la competitividad y asegurar un futuro más sostenible.

Según [4] las ciudades inteligentes utilizan y aplican las nuevas tecnologías en la administración de los servicios urbanos: Servicios sociales, distribución de energía y gobernanza.

En [5] se define una ciudad inteligente como una ciudad innovadora que utiliza Tecnologías de la Información y Comunicación (TIC) y otros medios para mejorar la toma de decisiones, la eficiencia de las operaciones, los servicios urbanos y su competitividad; a la vez que se garantiza la atención de las necesidades a las generaciones actuales y futuras en relación con los aspectos económicos, sociales y medioambientales.

B. Características y Componentes

En la Tabla 1 se relacionan las características y componentes propios de aplicaciones para ciudades inteligentes, de acuerdo al análisis de un conjunto de sistemas desarrollados.

Tabla 1. Características y componentes de las aplicaciones para ciudades inteligentes

Sistema	Componentes	Características
Parqueaderos inteligentes (<i>Smart Parking</i>) [6]	<ul style="list-style-type: none"> - Sensores de proximidad inductivos - Detector de metales - <i>Loop</i> de piso - RFID - Pulsadores - Placa electrónica Arduino - Reloj en tiempo real - Wifi - PLC (Programmable Logic Controllers) - LED - Sensores de presencia - Circuito detector de peso - Sensores de tope - Semáforos 	<ul style="list-style-type: none"> - Infraestructura de comunicaciones fija y móvil - Interfaz natural e intuitiva - Sensible al contexto - Fiabilidad y seguridad - Multimodal - Ubicuidad - Procedimiento concurrente - Fiabilidad - Capacidad de evolución - Tiempo real - Adaptativo - Invisible - Embebido - Predictivo - Personalizado - Dispositivos invisibles - Infraestructura de comunicaciones fija y móvil - Redes dinámicas de dispositivos distribuidos - Interfaz natural e intuitiva - Fiabilidad y seguridad - Computación ubicua - Comunicación ubicua - Interfaces inteligentes
Señalización inteligente del tráfico [7]	<ul style="list-style-type: none"> - Cámaras SmartView - Redes de sensores inalámbricas - CHIP - Raspberry Pi - Bluetooth - ZigBee - Wifi - Ethernet - GPRS/3G - Middleware Servicios Web - API REST - Magnetómetro para detección de vehículos - Semáforos - Paneles informativos 	
Iluminación inteligente de las calles [8]	<ul style="list-style-type: none"> - Luminarias LED - Sistema de conversión AC/AC - Optoacoplador - Sensor de luz - Wifi - Plataforma Arduino - Arduino Shield Ethernet - Arduino Shield Wi-fi - Arduino Shield Yun - Drivers para módulos LED - Pulsadores - Sensores o detectores - Unidades de control - Repetidores - Herramientas de configuración y de monitorización 	
Detección de disparos [9]	<ul style="list-style-type: none"> - Sensores acústicos - Máquina de clasificación - Notificaciones push - Micrófonos - Alarmas - Cámaras 	

Sistema	Componentes	Características
Calidad del aire/partículas en aire [10]	<ul style="list-style-type: none"> - Transmisores de temperatura - Conjunto integrado de sensores: Pluviómetro, anemómetro, temperatura, humedad, radiación solar y UV - Sensor Dräger X-am 5000 - PM10-Bombas para material particulado - Equipos de posicionamiento – GPS - Purificadores de aire - Sensores inteligentes para monitoreo de calidad del aire - Sensor de polvo - LED infrarrojo y un fototransistor - Sensores de CO2 - Sensores de Telaire - Tecnología infrarroja no dispersiva (NDIR) - Fuente infrarroja - Guía de ondas - Filtros infrarrojos - Detector termopila micromecanizado 	<ul style="list-style-type: none"> - Infraestructura de comunicaciones fija y móvil - Interfaz natural e intuitiva - Sensible al contexto - Fiabilidad y seguridad - Multimodal - Ubicuidad - Procedimiento concurrente - Fiabilidad - Capacidad de evolución - Tiempo real - Adaptativo - Invisible - Embebido - Predictivo - Personalizado - Dispositivos invisibles
Gestión energética [11]	<ul style="list-style-type: none"> Sensores con el sistema LCN (Red de Control local) Sensor de temperatura y receptor infrarrojo Sensor de humedad Detector de movimiento y sensor de luminosidad Detector de dióxido de carbono Sensor de luminosidad exterior Sensores para la gestión de la energía Sensores de luz Detectores de movimiento y de presencia Sensores de temperatura 	<ul style="list-style-type: none"> - Infraestructura de comunicaciones fija y móvil - Redes dinámicas de dispositivos distribuidos - Interfaz natural e intuitiva - Fiabilidad y seguridad - Computación ubicua - Comunicación ubicua - Interfaces inteligentes

Fuente: Elaboración propia.

C. Metodologías de Ingeniería de Requisitos Aplicables

En [12] se propone una metodología de requisitos de software para aplicaciones de *AmI* (*Ambient Intelligence*), la cual sugiere una fase de modelado de negocios, una fase de modelado de tecnología, una fase de modelado de interacción y una fase general de modelado, y en esta última fase se proporciona información acerca de las metas que se desea lograr.

Los autores en [13] proponen la metodología de especificación *Agile Requirements de SCRUM*, aplicable a soluciones móviles, la cual plantea un

documento de *back log* que define las diferentes funcionalidades y requerimientos del producto que desea el *stakeholder*. Una vez se reúnen estas funcionalidades del producto, se inicia la fase de priorización.

En [14] se presenta una metodología denominada DoRCU (Documentación de Requerimientos Centrada en el Usuario), caracterizada por su flexibilidad y orientación al usuario, la cual se apoya en diversos métodos, técnicas y herramientas ya desarrollados por otros autores, pero sin comprometerse con las tendencias de un modelo en particular.

Una metodología para la ingeniería de requisitos de sistemas ubicuos - ABC-Besoins, es la propuesta en [15], la cual está orientada por metas y agentes y proporciona formalismos para representar el conocimiento básico de un dominio y proceder a identificar los requisitos; brindando continuidad en el proceso de Ingeniería de Requisitos (IR) desde la elicitación de requisitos hasta la generación de un modelo conceptual y de un modelo de diseño.

En [16] se propone la ingeniería de requisitos para ambientes inteligentes (R4IE), la cual establece objetivos de alto nivel, determina el ámbito de trabajo, identifica las partes interesadas, las tareas y funcionalidad y las cualidades de rendimiento del sistema, y determina los perfiles del interesado.

En la Tabla 2 se presenta un resumen de las metodologías analizadas con respecto a la guía técnica presentada en el capítulo, teniendo en cuenta las características propias de cada metodología en relación con el desarrollo de aplicaciones para

ciudades inteligentes. Se destaca el valor agregado de la propuesta presentada con respecto a la especificación de requisitos. Las características que se comparan en la Tabla 2 son las siguientes.

Uso de formatos: La Utilización de formatos o plantillas para la especificación de requisitos.

Interacción humana – tecnología: la capacidad de la tecnología de dar respuesta a impulsos dados por los humanos y que deben reflejarse en los requisitos.

Interacción tecnología – tecnología: la capacidad de responder a solicitudes dadas entre tecnología y tecnología para llevar a cabo una tarea determinada.

Componentes tecnológicos: la participación de componentes tecnológicos necesarios en la definición de los requisitos.

Características de las ciudades inteligentes: Las metodologías contemplan particularidades explícitas de las ciudades inteligentes.

Tabla 2. Resumen de las metodologías analizadas con respecto a la guía técnica presentada en el capítulo

	Uso de formatos	Interacción humano – tecnología	Interacción tecnología – tecnología	Componentes tecnológicos	Características de las ciudades inteligentes
Metodología para modelar la inteligencia ambiental aplicaciones que usan i * framework.	No	Sí	Sí	Sí	Sí
Metodología de Especificación Agile Requirements de SCRUM	Sí	No	No	No	No
Metodología DoRCU para la Ingeniería de Requerimientos	Sí	No	No	No	No
Ingeniería de requisitos para Ambientes Inteligentes (R4IE).	Sí	Sí	Sí	Sí	Sí
Modelo de contexto y de dominio para la ingeniería de requisitos de sistemas ubicuos	Sí	No	No	Sí	No
Guía técnica para la especificación de requisitos para aplicaciones en ciudades inteligentes (propuesta)	Sí	Sí	Sí	Sí	Sí

Fuente: Elaboración propia.

III. RESULTADOS

A continuación, se presenta la guía técnica para la especificación de requisitos de software en ciudades inteligentes, que recoge las buenas prácticas de algunas metodologías y los formatos aplicables. Todo esto con base en la revisión y el análisis realizados en las sesiones anteriores con respecto al concepto de ciudades inteligentes, sus componentes y características. En la figura 1 se presenta la visión general de la guía técnica.



Fig. 1. Visión general de la guía técnica

A. Fase 1: Definición del Entorno

Denomina al conjunto de características o normas que regulan, describen y definen el funcionamiento de los lugares en los cuales funcionará el sistema que se va a desarrollar (parqueaderos, centros comerciales, hogares, empresas, negocios, etc.). El entorno se caracteriza por un conjunto de factores que delimitan el ámbito en el que estos lugares actúan, y establece las condiciones en que estos se tienen que desenvolver. Los factores son: económicos, sociales, tecnológicos, ambientales, de producción, de estructura organizacional y de recursos humanos.

B. Fase 2: Entorno Tecnológico Actual

Describe todos los elementos tecnológicos que están presentes en el entorno y que son relevantes para su funcionamiento (máquinas, herramientas, servicios, sensores, computadores, Smartphones, impresoras, cámaras de TV, etc.). Estos elementos son primordiales para reconocer la capacidad inteligente del entorno.

Hardware: Descripción de los componentes tecnológicos que están disponibles en el entorno en que se desarrollará la aplicación, y que ésta puede utilizar, tales como circuitos de cables y circuitos de luz, placas, actuadores, cámaras SmartView, controladores, sensores, etc.

Software: Se describen los programas informáticos que están instalados y disponibles en el entorno en el cual se desarrollará la aplicación.

C. Fase 3: Necesidades del Entorno

Describe las carencias y posibles mejoras que puede tener el entorno y que impiden desempeñar adecuadamente las funciones que se llevan a cabo. Se describen también los procesos que se pretende optimizar y darles solución, al igual que se busca el aprovechamiento de factores ambientales, energéticos, espaciales, entre otros. Esto con el fin de crear bienestar entre las personas que interactúan en el entorno.

D. Fase 4: Funcionalidad para Implementar en el Entorno

Describe la funcionalidad general que tendrá el software para solucionar la necesidad identificada. La descripción de la función es importante para detectar posibles necesidades que se tendrán en cuenta en la especificación de requisitos para un correcto desarrollo del software.

E. Fase 5: Especificación de Requisitos

Para la especificación de requisitos se tendrá en cuenta el siguiente proceso:

Identificación del requisito: Es un código alfanumérico que identifica al requisito.

Nombre del requisito: Denomina el título del requisito.

Componente: Lista de elementos eléctricos, electrónicos, electromecánicos y mecánicos con los cuales está asociada la funcionalidad del requisito.

Característica asociada: Se define la cualidad del requisito. Algunas cualidades son:

- **Embebido:** Dispositivos integrados en otros productos, y que a su vez controlan determinadas funciones en un sistema de computación en tiempo real.
- **Natural:** Ausencia de interacción con los sistemas informáticos.
- **Invisible:** Los dispositivos se deben integrar en el entorno de manera natural y discreta.
- **Ubicua:** Los dispositivos están presentes en cualquier parte del entorno inteligente que se va a desarrollar, deben interactuar entre sí y con el usuario, permitiendo al sistema acceder a la información que estos ofrecen desde cualquier lugar, mediante la implementación de redes inalámbricas.

Descripción del requisito: Exposición en detalle de la necesidad que se va a solucionar con el software, debe ser adecuada y precisa. Es importante que esta descripción contemple los siguientes elementos: Función que va a cumplir; objeto sobre el que se aplicará la función; acción asociada al componente inteligente empleado. Por ejemplo: Encender/apagar las lámparas automáticamente al detectar movimiento vehicular o humano.

Función	Objeto	Función asociada
Encender / apagar	Las lámparas	Detectar movimiento

Prioridad: Se define el nivel de importancia del requisito, de acuerdo a factores como aspectos técnicos, de recursos humanos, etc. Se asigna un valor numérico o un valor enumerado: prioridad alta, media y baja.

Restricciones: Indicar las limitaciones para tener en cuenta en el diseño y/o desarrollo de la solución asociadas al requisito, tales como: normas aplicables, estándares relacionados, operación, implementación, etc.

Interacción humano - tecnología: El usuario interactúa directamente con los componentes tecnológicos asociados a la solución del requisito de la aplicación que se va a desarrollar.

Interacción tecnología - tecnología: Interacción de los componentes tecnológicos sin la intervención humana.

En la Tablas 3 y 4 se muestran las plantillas aplicables.

En la Tablas 5 y 6 se muestra un ejemplo de aplicación de las plantillas en un parqueadero inteligente.

Tabla 3. Plantilla para la definición del entorno

FASE I	Definición del entorno
FASE II	Descripción del entorno tecnológico actual
Hardware	
Software	
FASE III	Establecer las necesidades del entorno
FASE IV	Descripción de la funcionalidad para implementar en el entorno

Fuente: Elaboración propia.

Tabla 4. Plantilla para la especificación de requisitos

Especificación de Requerimientos Funcionales (ERS)	
Identificación del requisito	
Nombre del requisito	
Componente	
Característica asociada	
Descripción del requisito:	
Características	
Prioridad	<input type="checkbox"/> Alta/Eencial <input type="checkbox"/> Media/Deseado <input type="checkbox"/> Baja/ Opcional
Restricciones	
Interacción humano - tecnología	<input type="checkbox"/> Sí <input type="checkbox"/> NO
Interacción tecnología - tecnología	<input type="checkbox"/> Sí <input type="checkbox"/> NO

Fuente: Elaboración propia.

Tabla 5. Ejemplo de definición del entorno

FASES PARA LA DEFINICIÓN DE REQUISITOS	
FASE I	Definición del entorno
Un parqueadero es un espacio público/privado destinado a la prestación del servicio de estacionamiento y cuidado de vehículos (automóviles y motos); el servicio puede ser gratuito o con fines comerciales mediante arrendamiento. Los parqueaderos poseen un horario definido de prestación del servicio, y se rigen por normas de tránsito y señalización. En el momento en que el usuario desee solicitar el servicio de parqueadero, el empleado de estacionamiento verifica la disponibilidad de espacios, si existen espacios disponibles se le solicita al usuario la información pertinente que le identifique como dueño del vehículo y se le asigna el espacio para ocupar. Cuando el usuario retira el vehículo del parqueadero, se le estima el tiempo de permanencia y de acuerdo a este se le cobra una tarifa ya establecida. En caso de estacionar su vehículo de manera permanente, el usuario puede acordar contratos con el parqueadero mediante el cobro de mensualidades.	
FASE II	Descripción del entorno tecnológico actual
Hardware	
Computador, cámaras de vigilancia, monitores de video, DVR (grabador de video digital), rótulos indicativos para mostrar la cantidad de espacios disponibles.	
Software	
Aplicativo de escritorio para administración del parqueadero, software del sistema de circuito cerrado de televisión, Office, navegadores web, software contable.	

FASES PARA LA DEFINICIÓN DE REQUISITOS	
FASE III	Establecer las necesidades del entorno
Se busca controlar el ingreso y la salida del vehículo (moto, automóvil) del parqueadero mediante el reconocimiento y almacenamiento de la placa en el sistema. Cuando el vehículo abandone el parqueadero es necesario calcular su tiempo de estadía, como también, detectar la presencia de los vehículos en los puestos de estacionamiento mostrando información en tiempo real de cuáles puestos están ocupados y cuáles no y cuántos son. También se pretende un control energético en el parqueadero.	
FASE IV	Descripción de la funcionalidad para implementar en el entorno
Sistema automático que permita controlar el ingreso y la salida del vehículo (moto, automóvil) del parqueadero, al momento de su ingreso el sistema debe reconocer y guardar la placa del vehículo al acercarse al parqueadero; toma el registro de entrada y, en caso de abandonar el parqueadero, el registro de salida, lo que permite calcular el tiempo de estadía y generar el valor de estacionamiento de acuerdo a ese tiempo. El sistema debe detectar la presencia de vehículos en un área específica y mostrar la información del área, como la información en tiempo real de cuáles y cuántos puestos de estacionamiento están ocupados, con la respectiva placa del vehículo y el tiempo de estadía, e identificar los puestos que están desocupados.	
El sistema debe emitir alertas en caso de detectar presencia de humo (incendios), y mostrar el lugar en pantalla donde ocurra el siniestro y activar automáticamente el sistema de extinción de incendios. De igual manera, debe emitir avisos inmediatos en caso de que se detecten fugas de agua o inundaciones, y mostrar el lugar en pantalla donde ocurra el suceso. Facilitar el ahorro de energía permitiendo que las lámparas se enciendan de noche y se apaguen de día de manera automática, y que se enciendan cuando alguien entre al parqueadero y se apaguen cuando salga.	

Fuente: Elaboración propia.

Tabla 6. Ejemplo de especificación de requisitos

FASE V. ERS – Especificación de Requerimientos Funcionales	
Id. del requisito	RF01
Nombre del Requisito	Control de ingreso y salida del vehículo
Componente	Sensor de parking, cámaras de vigilancia
Característica asociada	Embebido, invisible, ubicua
Descripción del requisito	Controlar el ingreso del vehículo en el parqueadero y su salida de éste mediante el reconocimiento de la placa
Características	Los usuarios deberán acercar su vehículo a un punto donde el sistema pueda reconocer y detectar la placa, con el fin de tomar el registro correspondiente y permitir el ingreso y su salida cuando el usuario desee abandonar el parqueadero
Prioridad	<input checked="" type="checkbox"/> Alta/Esencial <input type="checkbox"/> Media/Deseado <input type="checkbox"/> Baja/Opcional
Restricciones	<ul style="list-style-type: none"> El sistema estará diseñado para controlar el ingreso y la salida de automóviles y motos, mas no de bicicletas u otros medios de transporte.
Interacción humano - tecnología	SÍ <input type="checkbox"/> NO <input checked="" type="checkbox"/>
Interacción tecnología - tecnología	SÍ <input checked="" type="checkbox"/> NO <input type="checkbox"/>

FASE V. ERS – Especificación de Requerimientos Funcionales	
Id. del requisito	RF02
Nombre del Requisito	Verificación de disponibilidad
Componente	Sensores de ultrasonido, módulo de comunicaciones, PC, microcontrolador central, módulo Ethernet, LED (indica la disponibilidad del espacio), sistema de guía de parqueo PGS, rótulos indicativos para mostrar a los usuarios la cantidad de espacios disponibles y la dirección de dicho espacio
Característica asociada	Embebido, invisible, ubicua
Descripción del requisito	Informar cuando un puesto de estacionamiento esté ocupado o desocupado
Características	El sistema podrá verificar la disponibilidad de los puestos de estacionamiento mediante rótulos indicativos, que le muestran al usuario los puestos de estacionamiento que estén disponibles y la dirección en que se encuentran dichos puestos; adicional a esto, cada uno de los puestos tiene un bombillo LED que advierte su disponibilidad: Rojo: ocupado; Verde: disponible
Prioridad	<input type="checkbox"/> Alta/Esencial <input checked="" type="checkbox"/> Media/Deseado <input type="checkbox"/> Baja/Opcional
Restricciones	<ul style="list-style-type: none"> En caso de que haya más de un puesto de estacionamiento disponible, el usuario podrá seleccionar el puesto que desee, mas el sistema no lo asignará automáticamente
Interacción humano - tecnología	SÍ <input type="checkbox"/> NO <input checked="" type="checkbox"/>
Interacción tecnología - tecnología	SÍ <input checked="" type="checkbox"/> NO <input type="checkbox"/>

Fuente: Elaboración propia.

IV. CONCLUSIONES

En este capítulo se hace un acercamiento conceptual al término de *ciudad inteligente* desde varios autores, donde se da a entender que existe un factor en común: son ciudades que interactúan con las Tecnologías de la Información y la Comunicación (TIC).

Se logra evidenciar, por medio de una revisión de literatura, que no existen metodologías propias para la especificación de requisitos de software para aplicaciones en ciudades inteligentes.

La guía técnica propuesta incorpora un conjunto de elementos y características propios de las

aplicaciones para ciudades inteligentes, que no se contemplan en metodologías de ingeniería de requisitos actuales.

La guía técnica permite realizar una contextualización general del dominio donde se va a desarrollar la aplicación y luego detallar cada uno de los requisitos que le darán la funcionalidad al sistema.

La guía técnica propuesta busca reconocer el entorno tecnológico actual del dominio, así como validar la interacción (humano - tecnología y tecnología - tecnología) que implica la funcionalidad del requisito.

REFERENCIAS

- [1] R. P. Dameri and C. Rosenthal-Sabroux, Eds., *Smart City. How to Create Public and Economic Value with High Technology in Urban Space*. Springer International Publishing Switzerland, 2014.
- [2] Fundación Universia, “Smart Cities: estas son las ciudades más tecnológicas del mundo”, España, 11 de agosto 2017. [En línea]. Disponible en: <http://noticias.universia.es/ciencia-tecnologia/noticia/2017/08/11/1154968/smart-cities-ciudades-tecnologicas-mundo.html>
- [3] E. Ontiveros, D. Vizcaíno y V. López Sabater, *Las ciudades del futuro: inteligentes, digitales y sostenibles*. Madrid: Ed. Ariel, Fundación Telefónica, 2016.
- [4] A. Preukschat, Smart cities. 2018. Disponible en: <http://libroblockchain.com/smart-cities/>.
- [5] M. Bouskela, M. Casseb, S. Bassi, C. De Luca y M. Facchina, *La ruta hacia las Smart cities: Migrando de una gestión tradicional a la ciudad inteligente*. Inter-American Development Bank, 2016.
- [6] M. Swatha and K. Pooja, “Smart car parking with monitoring system”, in *IEEE International Conference on System, Computation, Automation and Networking (ICSCA 2018)*, pp. 1-5.
- [7] C. Izquierdo Martínez, “Aplicación de la IoT al ámbito del transporte. Auto-gestión del tráfico de vehículos inteligentes”, tesis de doctorado, Universitat Politècnica de València, 2017.
- [8] R. Corvalán, N. Sanabria, E. Ferrari, V. Titiosky, A. Amarilla, A. Cuevas, V. Sabaj y H. Fleitas, “Aplicación de criterios de optimización energética y seguridad en la iluminación y confort, en calles y avenidas”, *Extensionismo, Innovación y Transferencia Tecnológica. Claves para el desarrollo*, vol. 5, pp. 177-186, 2019.
- [9] D. S. Lawrence, N. G. La Vigne, M. Go and P. S. Thompson, “Lessons learned implementing Gunshot Detection Technology: Results of a process evaluation in three major cities”, *Justice Eval. J.*, No. 1, pp. 109-129, 2018.
- [10] M. Penza, D. Suriano, M. Villani, L. Spinelle and M. Gerboles, “Towards air quality indices in smart cities by calibrated low-cost sensors applied to networks”, in *Proceedings of the IEEE Sensors 2014*, Valencia, Spain, 2-5 November 2014, pp. 2012-2017.
- [11] L. Benítez y M. Ortega, “Las TIC y la gestión de los desafíos de sostenibilidad energética de las ciudades inteligentes”, *Economía Industrial*, N.º 395, pp. 87-94, 2015.
- [12] A. Guzman, A. Martínez, F. V. Agudelo, H. Estrada-Esquivel, J. P. Ortega and J. Ortiz, “A methodology for modeling ambient intelligence applications using i* framework”, in *Proceedings of the Ninth International i* Workshop co-located with 24th International Conference on Requirements Engineering (RE 2016)*, L. López y Y. Yu, eds., vol. 1674 of CEUR Workshop Proceedings, Beijing, China, September 12-13, 2016, CEUR-WS.org, pp. 61-66.
- [13] K. Vlaanderen, S. Jansen, S. Brinkkemper and E. Jaspers, “The agile requirements refinery: Applying SCRUM principles to software product management”, *Information and Software Technology*, vol. 53, pp. 58-70, 2011.
- [14] M. G. Báez y S. I. Barba Brunner, “Metodología DoRCU para la Ingeniería de requerimientos”, in *Workshop in Engineering of Requirements (WER)*. 2001, pp. 210-222.
- [15] L. González y G. Urrego, “Modelo de contexto y de dominio para la ingeniería de requisitos de sistemas ubicuos”, *Revista Ingenierías Universidad de Medellín*, vol. 9, N.º 17, pp. 151-167, 2010.
- [16] C. Evans, L. Brodie and J. Augusto, “Requirements Engineering for Intelligent Environments”, in *Proceedings of the 10th International Conference on Intelligent Environments (IE'14)*, 2014, pp. 154-161. IEEE Press. Disponible en: https://www.researchgate.net/publication/286679790_Requirements_Engineering_for_Intelligent_Environments

Capítulo III

Pruebas Automáticas para Evaluar Cursos de Programación de Computadores

Milton de Jesús Vera Contreras - miltonjesusvc@ufps.edu.co

Universidad Francisco de Paula Santander (UFPS) Cúcuta

Matías Herrera Cáceres - matiashc@ufps.edu.co

Docente-investigador, Universidad Francisco de Paula Santander (UFPS) Cúcuta

Óscar Alberto Gallardo Pére - oscargallardo@ufps.edu.co

Docente-directivo, Universidad Francisco de Paula Santander (UFPS) Cúcuta

I. INTRODUCCIÓN

En el año 2011, Marc Andreessen pronunció la popular frase “el software se está comiendo el mundo” (en inglés, “software is eating the world”) [1], y desde entonces es una realidad que no se ha detenido, pues el software es el núcleo de las tecnologías de la información y la comunicación (TIC) que hoy protagonizan la cuarta revolución industrial o Industria 4.0 [2]. En contraste, en el contexto colombiano y latinoamericano hay un déficit y existen debilidades en materia de talento digital [3], [4], por lo que cobra especial importancia la formación en programación de computadores [5], un área en la que enfocan sus esfuerzos los empresarios y el gobierno [6].

Ahora bien, históricamente la formación en programación de computadores ha sido un problema educativo complejo [6]-[13]. Por lo tanto, al considerar el panorama descrito sobre el talento digital, es de gran interés y alto impacto cualquier esfuerzo por facilitar, agilizar y mejorar la educación en esa área. En ese orden de ideas, el programa de Ingeniería de Sistemas de la Universidad Francisco de Paula Santander de Cúcuta (UFPS), en su ejercicio semestral de autoevaluación, realizó en el año 2017 un análisis detallado de las calificaciones en las asignaturas de Programación de Computadores. Los dos principales hallazgos de dicho análisis fueron el alto nivel de reprobación y las bajas calificaciones de los estudiantes que aprobaban. De acuerdo al análisis, más allá de buscar

las causas, era urgente intervenir para mejorar la situación. En consecuencia, se formuló un proyecto de investigación-acción con el propósito de reducir la cantidad de estudiantes que reprobaban, mejorar el nivel de calificaciones e innovar con calidad en el aprendizaje y la enseñanza de la programación de computadores.

En el marco de dicho proyecto se priorizó la acción, para lo cual se propuso la iniciativa de experimentar usando pruebas automáticas como instrumento para la evaluación de los estudiantes. Dicha iniciativa se inspiró en tres tópicos *ad hoc*, conocidos previamente por los profesores involucrados en el proyecto:

- Las experiencias y lecciones aprendidas de la programación competitiva, en la cual los estudiantes se enfrentan a competencias de programación de computadores, cuyo éxito depende de la rapidez con la cual se resuelven varios problemas [14].
- La evaluación automática, que consiste en automatizar la retroalimentación que recibe el estudiante en sus tareas y evaluaciones de programación de computadores [15]-[18].
- La estrategia de aprendizaje experiencial o vivencial, en especial el aprendizaje basado en problemas y aprendizaje basado en retos [19], que enfatiza en aprender haciendo.

La estrategia educativa propuesta de usar pruebas automáticas se resume en que los

profesores diseñan pruebas para los ejercicios de programación de computadores con los que se evalúa a los estudiantes, dichas pruebas funcionan como una rúbrica de evaluación y son automáticas, no requieren intervención humana. En la medida en que un profesor o un equipo de profesores diseñen varios ejercicios y sus pruebas, se consigue con el tiempo un banco de ejercicios, lo cual propicia el entrenamiento autónomo de los estudiantes y la retroalimentación inmediata, sin dependencia del profesor. Además, se cambia la dedicación del tiempo del profesor: más tiempo diseñando ejercicios y pruebas y muy poco tiempo calificando y se elimina el sesgo en la evaluación, pues es una máquina la que evalúa según criterios objetivos planeados al diseñar las pruebas.

El propósito de este documento es divulgar los detalles y resultados de esta iniciativa, para lo cual el documento se organiza de la siguiente forma: En primer lugar, se expone el análisis realizado a las calificaciones en las asignaturas de programación de computadores. Luego se reseña la literatura revisada para soportar o descartar los tópicos *ad hoc* en los que se inspira la propuesta de pruebas automáticas. En tercer lugar, se describe en detalle la estrategia, las actividades de aprendizaje y herramientas tecnológicas. Posteriormente, se muestran los resultados obtenidos, contrastando el desempeño de los estudiantes con la evaluación que dichos estudiantes realizaron al profesor. Finalmente, se plantean las conclusiones, se discuten las implicaciones educativas y se proponen trabajos futuros.

II. ANÁLISIS DE LAS CALIFICACIONES EN ASIGNATURAS DE PROGRAMACIÓN DE COMPUTADORES (BEFORE TEST)

Al inicio de cada semestre académico, el programa de Ingeniería de Sistemas de la Universidad Francisco de Paula Santander (UFPS), sede Cúcuta, en cumplimiento de la política institucional de autoevaluación, revisa con el equipo de profesores las fortalezas y debilidades en cada una de las asignaturas, como mecanismo de retroalimentación que permite emprender acciones de mejoramiento. En el caso específico de

Programación de Computadores, el programa tiene cuatro asignaturas que se cursan en los primeros cuatro semestres: Fundamentos de Programación, Programación Orientada a Objetos I, Programación Orientada a Objetos II y Estructuras de Datos.

Para cada una de estas asignaturas se analizaron las calificaciones de todos los semestres durante el periodo comprendido entre el segundo semestre de 2012 y el segundo semestre de 2016, un total de 1.544 registros (Figuras 1 y 2), según el Sistema de Información Académica. En estos registros no se considera la deserción por cancelación de una asignatura antes de finalizar el semestre.

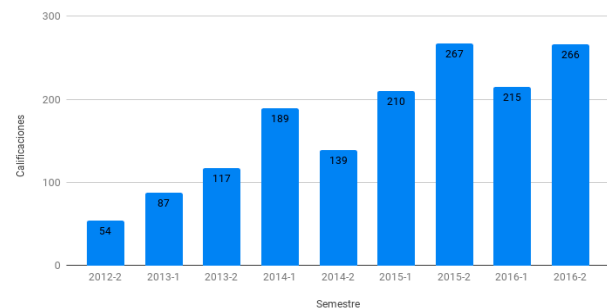


Fig. 1. Cantidad de registros de calificaciones por semestre

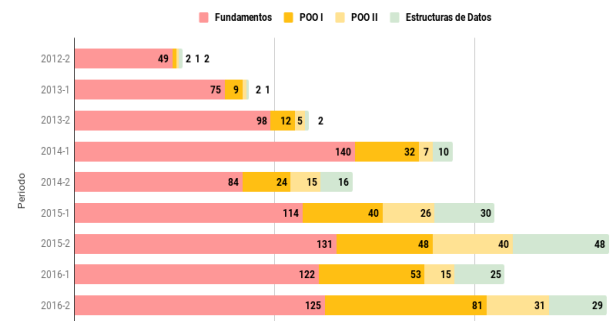


Fig. 2. Cantidad de registros de calificaciones por semestre y asignatura

Como se puede apreciar en las figuras, la cantidad de estudiantes se incrementa cada semestre, en lugar de mantenerse constante. Este comportamiento se explica por la retención de estudiantes que reprueban las asignaturas de Programación de Computadores. En consecuencia,

el análisis posterior se enfocó en tres preguntas: (i) ¿Cuál es el nivel de aprobación y reprobación de las asignaturas (se aprueba < 3.0)?, (ii) ¿Cuánto tarda un estudiante en aprobar una asignatura? (iii) ¿Cuál es el nivel de las calificaciones en estas asignaturas cuando el estudiante aprueba?

Para la primera pregunta, en promedio el 63 % de los estudiantes aprueban y el 37 % restante reprueba. El único curso que no cumple con este patrón de comportamiento es Fundamentos de Programación de primer semestre, el cual tiene una relación de 34 % aprobados y 66 % reprobados (Figura 3).

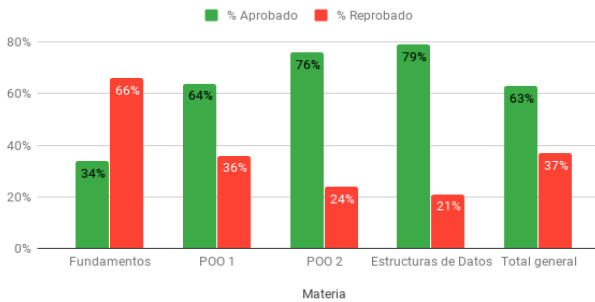


Fig. 3. Porcentaje de aprobación y reprobación por asignatura

Considerando la importancia del primer semestre, se analizó en detalle la asignatura Fundamentos de Programación. Se identificó que la situación mejoró en los últimos semestres y justo en el segundo semestre de 2016 se logró invertir ese comportamiento (Figura 4), con un 52 % aprobados y 48 % reprobados.

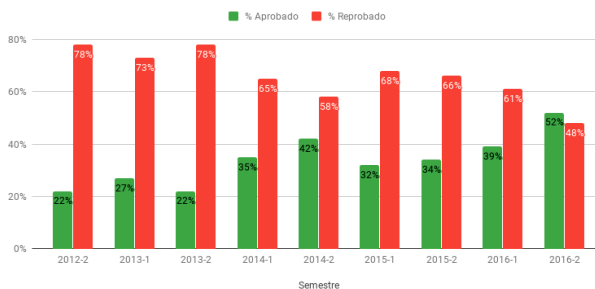


Fig. 4. Porcentaje de aprobación y reprobación en Fundamentos de Programación

La conclusión evidente para la primera pregunta es que hay una tasa alta de reprobación de las asignaturas de Programación de Computadores. Y la situación es más evidente al responder la segunda pregunta: Sólo nueve estudiantes aprobaron la primera vez una asignatura de Programación de Computadores, sin necesidad de repetirla, 234 estudiantes repitieron una vez, 101 repitieron dos veces, 50, tres veces y 39 estudiantes repitieron más de tres veces, con un máximo de nueve veces (Figura 5). Es importante aclarar que estas calificaciones incluyen todos los intentos que ofrece la UFPS para que un estudiante apruebe, como exámenes de habilitación y cursos en vacaciones.

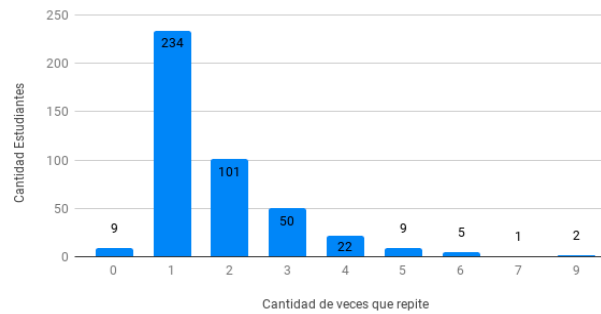


Fig. 5. Semestres requeridos para aprobar Programación de Computadores

En general, los 1.544 registros de calificaciones analizados correspondieron a 433 estudiantes y de esos registros 774 fueron de calificaciones reprobadas, lo cual arroja un promedio de 2,8 intentos para aprobar una asignatura de Programación de Computadores. Los otros 770 registros de calificaciones aprobadas se distribuyeron en cinco rangos para responder la tercera pregunta (Figura 6).

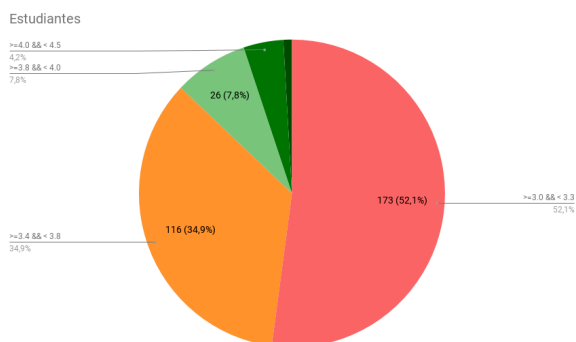


Fig. 6. Nivel de calificaciones aprobadas en asignaturas de Programación de Computadores

Siguiendo la Figura 6, el promedio de las calificaciones aprobadas es de 3,41, el porcentaje más alto (87,05 %) corresponde a calificaciones inferiores a 3,80 y sólo el 12,95 % obtiene calificaciones de nivel alto ($\geq 3,80$). La Figura 7 permite apreciar en detalle el promedio y la desviación estándar de las calificaciones aprobadas, reprobadas y el total, de cada una de las asignaturas. El promedio de las calificaciones reprobadas es 2,0 y el promedio total es 2,70.

Asignatura	Aprobado		Reprobado		Total	
	Promedio	Desviación	Promedio	Desviación	Promedio	Desviación
Fundamentos	3,37	0,39	1,93	0,64	2,44	0,89
POO 1	3,37	0,40	2,18	0,61	2,95	0,75
POO 2	3,51	0,44	2,38	0,43	3,28	0,63
Estructuras Datos	3,48	0,46	2,43	0,50	3,29	0,62
Total general	3,41	0,42	2,00	0,64	2,70	0,89

Fig. 7. Promedio y desviación estándar de calificaciones en asignaturas de Programación de Computadores

De acuerdo al análisis de las calificaciones, se derivaron tres conclusiones:

- Hay una tasa alta de reprobación de asignaturas de Programación de Computadores.
- Hay un bajo nivel en las calificaciones de los estudiantes que aprueban Programación de Computadores.
- Las dos asignaturas que deben intervenir son: Fundamentos de Programación y Programación Orientada a Objetos, de primero y segundo semestre, respectivamente.

Considerando las conclusiones, más allá de buscar las causas de la situación identificada, para el Programa de Ingeniería de Sistemas de la Universidad Francisco de Paula Santander (UFPS), Cúcuta, era fundamental intervenir para mejorar la situación. En ese sentido, se procedió a identificar en la literatura experiencias exitosas a fin de implementar acciones inmediatas para reducir la tasa de reprobación y mejorar el nivel de calificaciones, como se expone a continuación.

III. REVISIÓN DE LITERATURA SOBRE PROBLEMAS Y SOLUCIONES DE LA FORMACIÓN EN PROGRAMACIÓN DE COMPUTADORES

Se realizó una búsqueda en SCOPUS y ACM, usando las palabras programación (programming), aprender (learn) y enseñar (teach) y filtrando por los campos título (TITLE), resumen (ABSTRACT) y palabras clave (KEYWORD). Se encontró divulgación abundante y permanente desde los años sesenta del siglo pasado, en revistas tanto de ciencias de la educación como de ciencias de la computación, así como en revistas de tópicos mixtos y aplicados. En SCOPUS se encontraron menos registros (473) que en ACM (1.483), pero en ambos el interés por el tema es creciente (Figura 8). Asimismo, al usar otras opciones se obtienen muchos más resultados, pues existen publicaciones enfocadas solo en aprendizaje (*learn*) o solo en enseñanza (*teach*), incluso se encuentran publicaciones sin estas palabras y en su lugar hay términos como educación (*education*) entrenamiento (*training*), instrucción (*instruction*) y formación (*schooling*), entre otros.

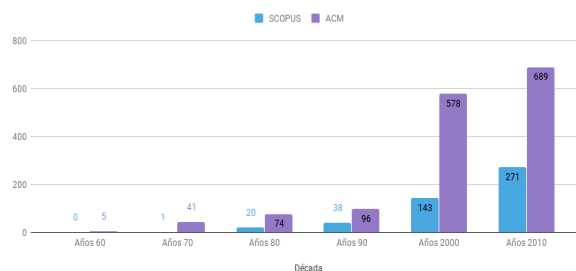


Fig. 8. Resultados de búsqueda en SCOPUS y ACM para “programming AND learn AND teach”

En general, la literatura sobre la formación en Programación de Computadores es abundante y ofrece diversos tópicos. Algunos autores se limitan a estudiar el problema sin especificar soluciones y otros abordan en detalle propuestas de solución. Para efectos de este documento es conveniente destacar los siguientes tópicos generales:

- Delimitación conceptual de qué es programación de computadores, estableciendo diferencias y relaciones con la resolución de problemas (*problem solving*), el pensamiento computacional (*computational thinking*), la codificación (*coding*) y la gerencia de proyectos de software (*software engineering*).
- Aspectos educativos diversos, como estilos de aprendizaje y docencia, aprender haciendo, conocimiento tácito y explícito, leer y escribir código, pedagogía y didáctica y la granularidad del currículo.
- Currículo del primer curso (*first course*) de programación de computadores, contenidos, actividades, evaluación, etc.
- Integración y articulación de la programación de computadores con otras áreas, como matemáticas, interfaces gráficas de usuario, juegos, redes, ingeniería del software, etc.
- Aspectos psicológicos y sociológicos como la motivación y la ansiedad o las preferencias y dificultades (del profesor y el estudiante) o el trabajo individual y grupal.
- El lenguaje y paradigma de programación usado para aprender a programar.
- El entorno de desarrollo IDE (*integrated development environment*) y sus características, como las facilidades de depuración (*debug, step-by-step*) y sugerencias de código (*code suggestion/completion*).
- El uso de herramientas tecnológicas de apoyo, como herramientas para visualización, diseño y automatización de tareas.
- El uso de herramientas para pruebas automáticas de software y evaluación automática de tareas de programación, que se deriva del tópico anterior.

Para la búsqueda sobre la idea específica de usar pruebas automáticas, se limitaron los resultados

al término evaluación formativa (*assessment*) y se omitieron otros términos (*test, evaluation, appraisal*), menos relacionados con el problema de interés. De manera consistente con la búsqueda anterior, en SCOPUS se encontraron menos registros (151) que en ACM (216), aunque en proporción diferente (Figura 9).

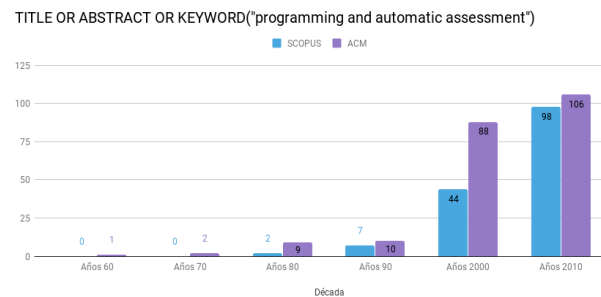


Fig. 9. Resultados de búsqueda en SCOPUS y ACM para “programming AND automatic assessment”

Puesto que el objetivo de la revisión de literatura era identificar soluciones y experiencias exitosas para abordar los problemas de la formación en programación de computadores, se realizó una lectura de los títulos y resúmenes de las publicaciones usando como criterio de ordenamiento la cantidad de citas, a fin de enfocarse en el impacto, y se profundizó en el detalle de las publicaciones que aportaban soluciones al contexto específico de la UFPS y que se relacionaban con los tópicos *ad hoc* que inspiraron la propuesta inicial, lo cual se expone a continuación.

A. Programación Competitiva

La programación competitiva es una actividad en la que equipos de tres estudiantes se enfrentan a resolver problemas complejos relacionados con matemáticas y ciencias de la computación, para lo cual deben escribir la mayor cantidad de programas de computador en el menor tiempo posible [20]. Este tipo de competencias iniciaron en 1974, y en la actualidad agremian mundialmente a muchos profesores y estudiantes en lo que se conoce como la Internacional Collegiate

Programming Contest (ICPC) [21]. En Colombia, la Asociación Colombiana de Ingenieros de Sistemas (ACIS) realiza desde 1986 la Maratón Nacional de Programación, que permite clasificar a la competencia latinoamericana, la cual otorga uno o dos cupos para la competencia mundial ICPC [22]. Asimismo, en Colombia existen dos ligas que mensualmente organizan competencias de entrenamiento: la Red de Programación Competitiva (RPC), que organiza competencias para toda Latinoamérica, desde México hasta Argentina [23], y la Colombian Collegiate Programming League (CCPL), que realiza competencias para estudiantes universitarios de Colombia [24].

El Programa de Ingeniería de Sistemas de la UFPS se vinculó y se ha mantenido vigente en este tipo de competencias desde hace cinco años, bajo el liderazgo de los mismos estudiantes, y ha logrado buenos resultados, en especial la motivación y el autoaprendizaje en tópicos que no son el eje del currículo del programa [14]. Por el éxito de esta experiencia, los estudiantes líderes del grupo desarrollaron un marco de trabajo y una plataforma de entrenamiento (*Training Center*), disponible en el sitio <http://trainingcenter.cloud.ufps.edu.co> [25]. En dicha plataforma se ha conformado un banco de problemas y de material de estudio que ha permitido a los estudiantes aprender y competir exitosamente a nivel nacional y latinoamericano.

Para el proyecto de Training Center se estudiaron diversas plataformas de programación competitiva, y se destacan las siguientes:

URI (<https://www.urionlinejudge.com.br/>)
UVA (<https://uva.onlinejudge.org/>)
Codeforces (<https://codeforces.com/>)
A20J (<https://a2oj.com/>)
Codechef (<https://www.codechef.com/>)
Boca (<http://bombonera.org/>)
Acepta el Reto (<https://www.aceptaelreto.com/>).

Todas estas plataformas permiten subir un programa de computador en varios lenguajes de programación (Java, C y C++, Python, C#, entre otros) y ofrecen retroalimentación inmediata, informando si el programa resuelve un problema

determinado y cumple con los requerimientos no funcionales de rendimiento en procesador y memoria RAM. Por ser un contexto de competencia, estas plataformas solo ofrecen retroalimentación total: indican si el programa cumple o no cumple, pero no hay pistas o indicaciones parciales para guiar al programador hacia una solución. No obstante, en el contexto del aprendizaje de programación de computadores, es deseable la retroalimentación parcial. En consecuencia, inspirados en el éxito de la programación competitiva, se propuso como una posible solución el uso de plataformas similares, pero para apoyar el aprendizaje de cursos de programación de computadores.

B. Pruebas Automáticas y VPL

El uso de pruebas automáticas en cursos de programación de computadores aparece en la literatura como evaluación automática (*automatic assessment*) y es algo que se ha probado con éxito desde hace cuatro décadas [15], [18], [26], [27]. Este tipo de plataformas abordan varios aspectos, y se destacan tres que aplican para el presente trabajo: En primer lugar, enfrentan el problema de la retroalimentación que necesita el estudiante cuando estudia sin presencia del profesor o tutor, con lo cual se propicia el entrenamiento autónomo e independiente, el autoaprendizaje y la motivación individual [12], [16], [18]. En segundo lugar, reducen la sobrecarga del profesor, puesto que para la evaluación de cada estudiante se requiere mucho tiempo [12], [15], [16]. Y tercero, permiten enfrentar el problema del plagio, puesto que se tiene un repositorio de problemas y soluciones que se pueden comparar de manera mucho más sencilla y rápida [28].

También, las plataformas de pruebas automáticas permiten usar varios lenguajes de programación, con lo cual los cursos se pueden orientar al desarrollo de habilidades, solución de problemas y conceptos fundamentales, en lugar de centrarse en aprender un lenguaje específico. Adicionalmente, el uso de pruebas automáticas permite introducir la técnica de Desarrollo de Software Dirigido por Pruebas (*Test Driven Software Development*), usada en el

contexto del software libre y los modelos ágiles, como los propuestos por Kent Beck [29] y Rober Martin [30], usados ampliamente en el mundo profesional de la Industria 4.0.

Para el presente trabajo se implementó la plataforma VPL (*Virtual Programming Lab*), una iniciativa de la Universidad Las Palmas de Gran Canaria (España) que se integra con el Sistema de Gestión de Aprendizaje Moodle, permite usar varios lenguajes de programación y ofrece verificación de plagio [18], [28]. VPL tiene licencia libre y abierta GNU/GPL, cuenta con documentación detallada, hay más de veinte publicaciones relacionadas en su sitio web [31] y las estadísticas de uso indican que 824 sitios Moodle usan VPL (Figura 10).

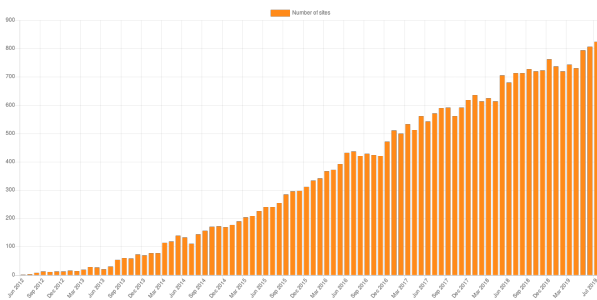


Fig. 10. Estadística de uso VPL en Moodle [31]

La arquitectura de VPL (Figura 11) oculta la complejidad del proceso de calificación, seguridad y lenguajes en una máquina virtual llamada “jaula de ejecución”. Dicha jaula recibe peticiones desde el servidor Moodle, lo cual se consigue mediante un componente (plugin). VPL ofrece además un editor de código fuente enriquecida que se ejecuta en el navegador (*browser*) usando HTML y JavaScript, en el cual el estudiante puede escribir su programa sin requerir un entorno de desarrollo (IDE). Este editor permite realizar exámenes presenciales y reduce la complejidad que involucra el aprendizaje de un IDE.

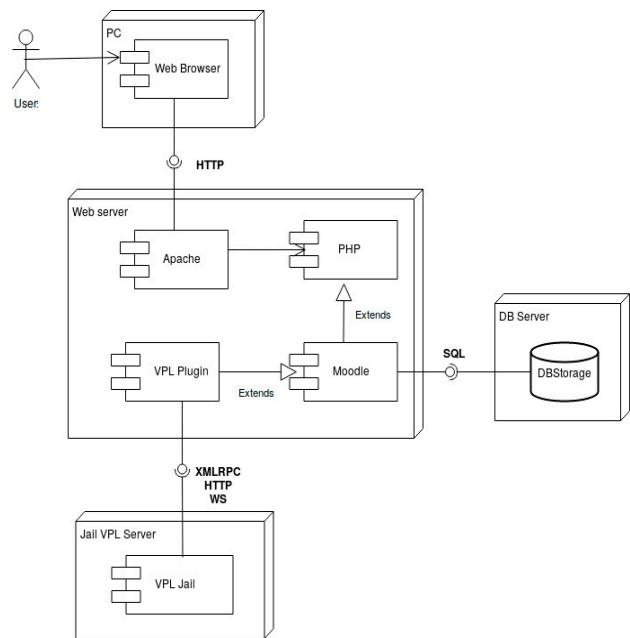


Fig. 11. Arquitectura de VPL y Moodle

Cada ejercicio de programación en VPL y Moodle tiene un enunciado y las pruebas. Existen dos alternativas para las pruebas: La primera son las pruebas de caja negra (Figura 12), que consisten en una lista de datos de entrada y salida llamados casos de prueba (*test cases*). VPL ejecuta el programa del estudiante, le suministra las entradas de prueba y luego compara las salidas del programa con las salidas de prueba. En cada caso de prueba se puede configurar que la calificación aumente o disminuya según el resultado y se pueden agregar comentarios de retroalimentación. Toda la información se almacena en un repositorio que puede ser consultado por el profesor y el estudiante, según corresponda. Este tipo de pruebas son ideales para cursos de primer semestre, donde se aprende el paradigma de programación estructurado/procedural.

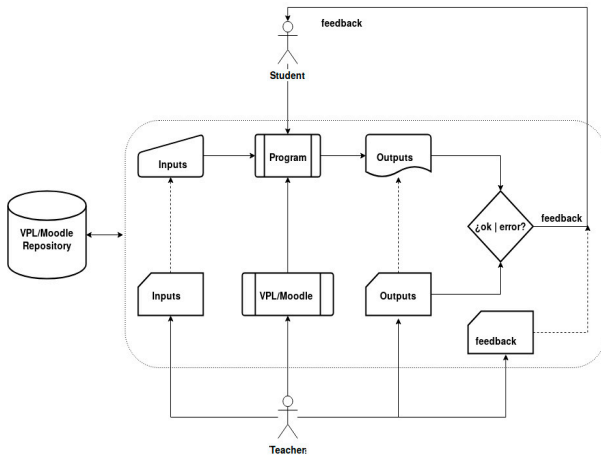


Fig. 12. Arquitectura de VPL y Moodle de pruebas de caja negra

La segunda alternativa de pruebas en VPL son las pruebas unitarias (Figura 13), usando librerías especializadas como JUnit.

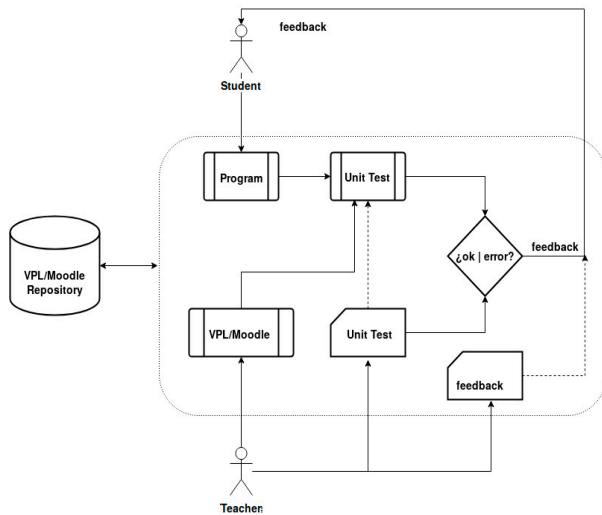


Fig. 13. Arquitectura de VPL y Moodle de pruebas unitarias

VPL ejecuta las pruebas unitarias usando como entrada el programa del estudiante. Las pruebas unitarias deben configurarse para que generen retroalimentación y calificación, para lo cual deben seguirse la documentación y los ejemplos en el sitio web de VPL. Este tipo de pruebas son ideales para cursos de segundo semestre, en los que se aprende el paradigma de programación orientado a objetos.

Para la calificación, VPL y Moodle permiten una escala porcentual de 0 a 100 o una escala específica seleccionada por el profesor. Además, tienen opciones de trabajo individual o en grupo y opciones para permitir varios intentos. Como cada ejercicio es una actividad de evaluación dentro del curso en Moodle, el profesor puede llevar un registro detallado de la dedicación y el desempeño del estudiante, y el estudiante puede intentar resolver el ejercicio tantas veces como lo permita el profesor.

Puesto que VPL es un plugin de Moodle, y junto a la jaula de ejecución se dispone del código fuente y la documentación con licencia GNU/GPL, los usuarios pueden personalizar el plugin y la jaula para que se adapten según las necesidades específicas. No obstante, las opciones que ofrece VPL para pruebas caja negra y unitarias son bastante sencillas de usar y hacen a VPL muy útil para implementar fácilmente la idea de pruebas automáticas en cursos de Programación de Computadores. Además, la integración con el Sistema de Gestión de Aprendizaje Moodle facilita su uso en contextos educativos virtuales, presenciales o mixtos, lo que permite aprovechar las ventajas tecnológicas para introducir estrategias educativas como el aprendizaje basado en retos.

C. Aprendizaje Basado en Retos

El aprendizaje basado en retos es un enfoque pedagógico centrado en el estudiante, que privilegia el hacer como medio para llegar al saber, con la particularidad de que el hacer se promueve estimulando al estudiante mediante un reto [19]. Un reto es una actividad de aprendizaje, tarea o evaluación, que consiste en una situación real o hipotética con la cual se desafía al estudiante a buscar el saber y a ponerlo en práctica, algo que se requiere diariamente en profesiones de ingeniería y ciencia [19], [32]. En el aprendizaje basado en retos se enfatiza la importancia de la creatividad y el ingenio, por eso el centro es el estudiante y no el contenido ni las actividades ni el profesor, el ser humano es el centro del aprendizaje. Esta idea se relaciona con la escuela de pensamiento de la psicología de

la programación de computadores, que plantea la importancia de abordar la programación como una actividad humana [7], [8], [33], [34]. Según la psicología de la programación de computadores cualquier persona puede aprender a programar y puede llegar a ser un experto programador, siempre y cuando se reciba una formación inicial apropiada, se tenga un entrenamiento permanente y se consiga una motivación apropiada y oportuna [8].

Una dificultad al implementar el aprendizaje basado en retos es la formulación de los retos y su posterior evaluación. Una vez un estudiante logra resolver un reto es fundamental suministrarle retos adicionales, y esto demanda una carga de trabajo para el profesor que es proporcional a la cantidad de estudiantes y que crece con el tiempo y de manera proporcional al aprendizaje [12], [35], [36]. En ese sentido, resulta de gran ayuda el uso de herramientas tecnológicas, como el enfoque descrito de pruebas automáticas con VPL, que permite disponer de un repositorio de retos y retroalimentación automática, y logra que el estudiante mantenga la motivación y además lidere su propio aprendizaje [15], [16], [18].

Como lo indica la literatura sobre los problemas y soluciones de la formación en programación de computadores, influyen mucho las prácticas educativas de ambos, profesores y estudiantes [6], [37], [38]. Asimismo, las experiencias exitosas se derivan de cambios innovadores en estas prácticas [11], [18], [39]. Es así que la literatura respalda la idea de introducir el aprendizaje basado en retos en combinación con pruebas automáticas, partiendo de la experiencia exitosa de la programación competitiva, como se explica a continuación.

IV. ESTRATEGIA, ACTIVIDADES DE APRENDIZAJE Y HERRAMIENTAS TECNOLÓGICAS

A. Curso de Primer Semestre – Fundamentos de Programación

La Figura 14 resume la estrategia implementada para primer semestre, cuyo objetivo es el aprendizaje de los Fundamentos de Programación.

En primer lugar se tienen dos momentos: el trabajo en la clase presencial, con el acompañamiento y supervisión del profesor y el trabajo independiente, ambos usando recursos tecnológicos. Siguiendo las ideas de la psicología de la programación, el aprendizaje de la programación de computadores se relaciona con leer y escribir código fuente [8]. En consecuencia, la estrategia prioriza leer código fuente en clase presencial con acompañamiento del profesor, y se deja como trabajo independiente escribir código, para lo cual el estudiante puede apoyarse en la retroalimentación de las pruebas automáticas.

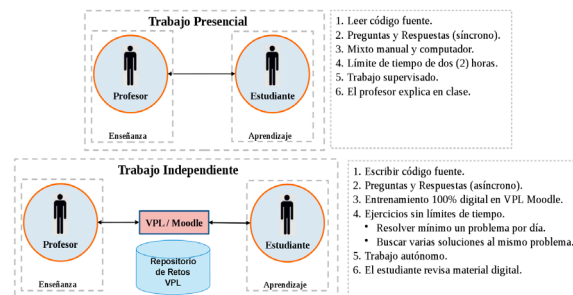


Fig. 14. Estrategia

En clase presencial se explican ejercicios de ejemplo y se resuelven dudas, siempre bajo la supervisión del profesor y con un tiempo límite de dos horas, que es el tiempo programado para cada una de las dos clases semanales. Este límite de tiempo implica que no todos tienen tiempo para preguntar y no todas las dudas se logran resolver. Hasta aquí no hay diferencias con un modelo clásico de clase de programación de computadores. La diferencia aparece al involucrar las pruebas automáticas, pues el estudiante, en su trabajo independiente, tendrá retroalimentación inmediata y sin límites de tiempo ni de intentos. La regla establecida es resolver como mínimo un problema por día, de lunes a viernes, e intentar encontrar varias soluciones diferentes al mismo problema.

Surge entonces un desafío para el profesor: lograr acopiar un banco de ejercicios suficiente para satisfacer la regla de resolver un problema diario. En la medida en que un profesor o un equipo de profesores diseñen varios ejercicios y sus pruebas, se consigue con el tiempo este banco de ejercicios.

La estrategia propicia el entrenamiento autónomo de los estudiantes y la retroalimentación inmediata sin dependencia del profesor. Ahora se cambia la dedicación del tiempo del profesor: más tiempo diseñando ejercicios y pruebas y muy poco tiempo calificando y, como ya se dijo, se elimina el sesgo en la evaluación, pues es una máquina la que evalúa según criterios objetivos planeados al diseñar las pruebas.

De acuerdo con esta estrategia, se implementó en Moodle el curso VPL (Virtual Programming Lab) Training, disponible bajo licencia Creative Commons (CC BY-NC-SA 4.0), con más de setenta problemas (<https://uvirtual.cloud.ufps.edu.co/course/view.php?name=VPLTraining>). El curso organiza los ejercicios en cinco unidades que abordan los tópicos principales de Fundamentos de Programación (Figura 15). El estudiante recibe en clase la explicación de cada tópico y debe desarrollar como trabajo independiente cada uno de los ejercicios, usando VPL y Moodle.

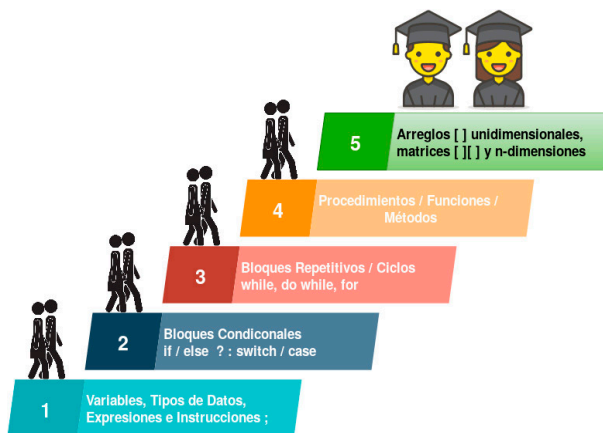


Fig. 15. Actividades de aprendizaje primer semestre

El estudiante puede hacer consultas por medios digitales (foro, chat y correo electrónico), las cuales se resuelven de manera personalizada. Asimismo, a través del foro y otros medios el estudiante puede compartir preguntas y soluciones con sus compañeros. En clase presencial el profesor enfoca

la explicación en los problemas más complejos, en las dudas que los estudiantes informaron y en los demás aspectos derivados de la revisión, que manualmente realiza el profesor como parte del acompañamiento y asesoría al curso. De este modo las actividades de aprendizaje se resumen en la resolución de problemas, la lectura y escritura de código fuente y una conversación permanente entre estudiantes y profesor sobre los tópicos y sus ejercicios, lo cual coincide mucho con el escenario real al desarrollar software en ámbitos profesionales.

En cuanto a las herramientas tecnológicas, además de VPL Moodle, se usa el entorno de desarrollo (IDE) BlueJ, por sus facilidades para visualización de conceptos y depuración [39], [40]. Los estudiantes pueden usar BlueJ en clase o desconectados, y en línea VPL les ofrece un entorno para programar directamente en el navegador y allí mismo conocer la retroalimentación (Figura 16). Por su parte, el profesor puede observar en VPL el detalle de todas las entregas e intentos de los estudiantes, hacer comentarios y evaluar el nivel de similitud o plagio (Figura 17).

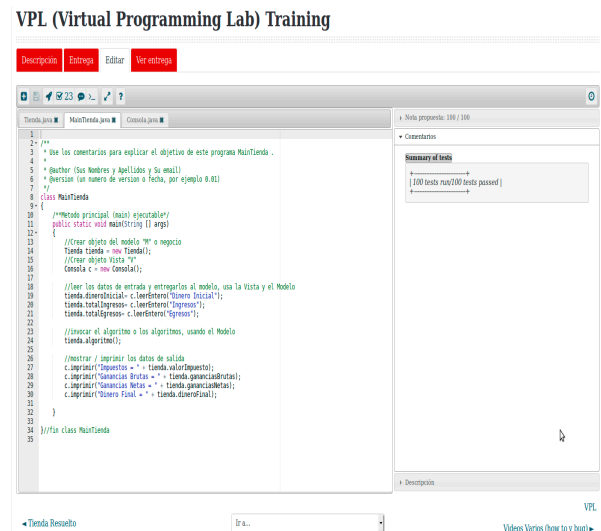


Fig. 16. Herramientas tecnológicas, entorno IDE de VPL Moodle para el estudiante

VPL (Virtual Programming Lab) Training						
Descripción						
Lista de entregas						
Simular						
Probar actividad						
Selección de entregas						
Todas las entregas						
Evaluación						
Nombre / Apellido(s)	Entregada el	Entregas	Calfificación	Evaluada por	Evaluada el	
1 Jose David Narajo Villan	Thursday, 5 de April de 2018, 21:07	4	100,00 / 100,00	Calfificación automática	Friday, 6 de April de 2018, 10:38	
2 Carlos Felipe González Galvis 1151750	Friday, 6 de April de 2018, 10:34	7	100,00 / 100,00	Calfificación automática	Thursday, 12 de April de 2018, 06:43	
3 MARIO ANTONIO ACOSTA ESPARZA	Friday, 6 de April de 2018, 10:36	6	100,00 / 100,00	Calfificación automática	Friday, 6 de April de 2018, 10:36	
4 ANGELY CAMILA CALDERON GARCIA 1151750	Friday, 6 de April de 2018, 10:36	2	100,00 / 100,00	Calfificación automática	Friday, 6 de April de 2018, 10:37	
5 JARLIN ANDRES FONSECA REBONON 1151750	Friday, 6 de April de 2018, 10:38	4	100,00 / 100,00	Calfificación automática	Friday, 6 de April de 2018, 10:38	
6 ANGELY NATALIA GARCIA BARBERA	Friday, 6 de April de 2018, 10:39	9	100,00 / 100,00	Calfificación automática	Friday, 6 de April de 2018, 10:40	
7 LEIDER MARTINEZ 1151706	Friday, 6 de April de 2018, 10:40	7	100,00 / 100,00	Calfificación automática	Friday, 6 de April de 2018, 10:40	

Fig. 17. Herramientas tecnológicas, entorno de VPL Moodle para el profesor

La implementación de esta estrategia, las actividades de aprendizaje y herramientas inició en el primer semestre del año 2017 y se ha mantenido desde entonces. En el segundo semestre del mismo año se avanzó a segundo semestre para darle continuidad al proceso, como se explica seguidamente.

B. Curso de Segundo Semestre – Programación Orientada a Objetos I

Para el segundo semestre se mantuvo la misma estrategia, ajustando las actividades de aprendizaje al Paradigma Orientado a Objetos y considerando los tópicos establecidos por el currículo del programa de Ingeniería de Sistemas (Figura 18).



Fig. 18. Actividades de aprendizaje segundo semestre

A diferencia de primer semestre, en segundo semestre las pruebas automáticas dejaron de ser caja negra y pasaron a ser pruebas unitarias, usando JUnit. En consecuencia, el estudiante puede ejecutar las pruebas unitarias sin conexión a VPL, usando únicamente BlueJ y publicando en VPL sólo la versión definitiva del ejercicio, lo cual dificulta hacer seguimiento al trabajo del estudiante. Una dificultad en segundo semestre corresponde a los estudiantes que no pasaron por el mismo proceso el semestre anterior, por lo que es necesario remitirlos al curso de VPL Training para nivelar. Algo que también conviene mencionar es que la cantidad de problemas se redujo a un promedio de veinte, considerando la densidad y complejidad del paradigma orientado a objetos, y a la fecha aún no se han acopiado en un único curso de entrenamiento como se hizo para el primer semestre.

Como se muestra a continuación, en la sección de resultados, durante dos años se ha mantenido en ambos semestres la estrategia, actividades de aprendizaje y herramientas tecnológicas, lógicamente con algunas pequeñas adaptaciones según la retroalimentación de los mismos estudiantes y la evaluación del profesor a cargo de la iniciativa.

V. RESULTADOS Y DISCUSIÓN (AFTER TEST)

En general los resultados se pueden considerar satisfactorios. Esta afirmación se fundamenta en las calificaciones finales obtenidas por los estudiantes, la tasa promedio de aprobación y la evaluación que los estudiantes realizan al profesor semestralmente. Las Figuras 19 y 20 permiten apreciar el porcentaje de estudiantes que aprobaron y reprobaron los cursos de primero y segundo semestres en los cuales se implementaron pruebas automáticas, con un promedio de 35 estudiantes en primer semestre y de 30 en segundo semestre.

% Aprobado y % Reprobado Primer Semestre

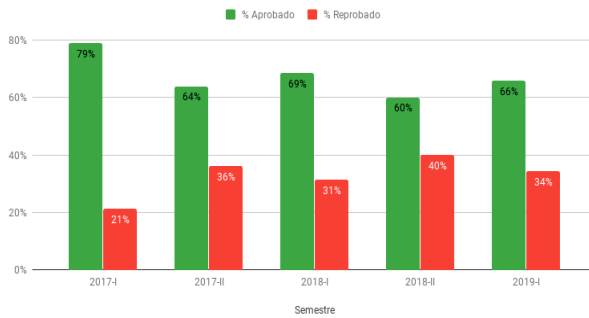


Fig. 19. Porcentaje de aprobación y reprobación primer semestre, usando pruebas automáticas

% Aprobado y % Reprobado Segundo Semestre

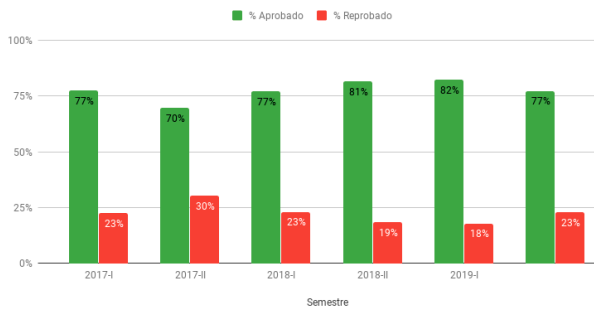


Fig. 20. Porcentaje de aprobación y reprobación segundo semestre, usando pruebas automáticas

Se realizó un análisis de los estudiantes que reprobaron en primer semestre y, en su mayoría, repiten una sola vez la asignatura. Durante los semestres analizados se ha conseguido mantener un nivel de reprobación aceptable y se observa una mejora en las calificaciones, con lo cual se ha minimizado la cantidad de estudiantes con calificaciones inferiores a 2,0, como se muestra en las Figuras 21 y 22.

Semestre	Aprobado		Reprobado		Total	
	Promedio	Desviación	Promedio	Desviación	Promedio	Desviación
2017-I	3,43	0,38	2,46	0,46	3,22	0,54
2017-II	3,39	0,36	2,59	0,51	2,94	0,61
2018-I	3,21	0,41	2,40	0,38	2,89	0,53
2018-II	3,33	0,43	2,12	0,38	2,76	0,67
2019-I	3,59	0,54	2,68	0,62	3,23	0,73
Total general	3,39	0,42	2,45	0,47	3,01	0,62

Fig. 21. Promedio y desviación de calificaciones en primer semestre, usando pruebas automáticas

Semestre	Aprobado		Reprobado		Total	
	Promedio	Desviación	Promedio	Desviación	Promedio	Desviación
2017-I	3,40	0,57	2,43	0,08	3,18	0,63
2017-II	3,29	0,37	2,67	0,15	2,97	0,42
2018-I	3,27	0,51	2,53	0,23	3,10	0,46
2018-II	3,09	0,48	2,42	0,32	2,96	0,47
2019-I	3,36	0,62	2,72	0,15	3,24	0,41
Total general	3,28	0,51	2,55	0,19	3,09	0,48

Fig. 22. Promedio y desviación de calificaciones en segundo semestre, usando pruebas automáticas

Por otra parte, en la evaluación docente que realizan los estudiantes cada semestre, el profesor a cargo del proyecto ha obtenido calificaciones y observaciones muy buenas (Figura 23), lo que indica que los estudiantes se sienten satisfechos con la metodología implementada.

Semestre	Primer Semestre	Segundo Semestre
2017-I	4,81	4,75
2017-II	4,85	4,75
2018-I	4,82	4,76
2018-II	4,81	4,68
2019-I	4,74	4,82
Promedio	4,81	4,75

Fig. 23. Calificaciones del profesor en la evaluación semestral que le hacen sus estudiantes

De acuerdo con las observaciones del profesor a cargo del proyecto, los estudiantes se mantienen motivados y en actividad permanente durante todo el semestre y se genera una dinámica de colaboración e interés por resolver los problemas en VPL Moodle. Una situación interesante es que algunos estudiantes tienden a cuestionar los resultados y argumentan que la plataforma o las pruebas están equivocadas. No obstante, una vez en clase se resuelven las inquietudes, los estudiantes aceptan los errores y retoman el ejercicio hasta resolverlo por completo. Asimismo, se experimentaron algunos inconvenientes de rendimiento, lo que generó la necesidad de migrar tanto Moodle como VPL a un modelo de Computación en Nube (*Cloud Computing*) usando una arquitectura de contenedores (*docker*), a fin de mejorar el rendimiento y facilitar el crecimiento y decrecimiento de recursos computacionales.

Finalmente, hay dos resultados adicionales que se han identificado durante el proyecto: en primer lugar, algunos estudiantes han comenzado desde primer semestre a utilizar plataformas de programación competitiva y a participar en el semillero, pues allí profundizan en tópicos que no se abordan en las clases y que les interesan y los motivan. En segundo lugar, en el curso de Programación Orientada a Objetos II, de tercer semestre, se realizó un experimento de introducir Programación Web y uso de Bases de Datos, considerando que la mayoría de los estudiantes provenían de los cursos donde se usaron pruebas automáticas. El experimento fue exitoso, y los estudiantes desarrollaron proyectos de aplicaciones web interesantes bajo un enfoque de innovación y emprendimiento.

VI. CONCLUSIONES Y TRABAJO FUTURO

Si bien falta mucho trabajo para generar conclusiones definitivas, los resultados obtenidos a la fecha indican que hay un nivel importante de satisfacción de los estudiantes y conviene continuar abordando el problema y mejorando el enfoque propuesto. Básicamente se mejoró la situación inicial, pero aún hay un volumen importante de estudiantes que reprueban y los niveles de calificación siguen siendo aceptables.

Es importante pasar ahora de la acción a la investigación, para identificar con mayor claridad las dificultades en el proceso de aprendizaje y las prácticas más apropiadas tanto de profesores como de estudiantes, a fin de aprobar exitosamente los cursos de Programación de Computadores y mejorar el nivel de calificaciones.

En el análisis presentado en este documento no se estudiaron en detalle los estudiantes que reprobaron, lo cual es algo fundamental para entender por qué es difícil aprender Programación de Computadores y cómo superar esa dificultad. De alguna manera, los estudiantes que reprueban demandan mayor atención por parte del profesor, y esto es necesario incorporarlo al enfoque propuesto a fin de reducir aún más la tasa de reprobación.

Aunque las calificaciones obtenidas por el profesor en la evaluación que le hacen sus estudiantes son buenas, no se puede afirmar que esto derive de la estrategia propuesta, las actividades de aprendizaje elaboradas o de la herramienta de pruebas automáticas empleada. Es necesario avanzar a una segunda etapa donde todos los profesores sigan el mismo enfoque y usen el mismo material educativo y tecnológico, a fin de identificar la influencia de la práctica docente específica.

Después de cinco semestres de trabajo, se tiene un curso de entrenamiento para estudiantes de primer semestre, con más de setenta problemas originales sobre tópicos de Fundamentos de Programación. El curso se liberó bajo licencia abierta Creative Commons (CC BY-NC-SA 4.0) y se puede acceder en la plataforma del programa de Ingeniería de Sistemas de la UFPS (<https://uvirtual.cloud.ufps.edu.co/course/view.php?name=VPLTraining>). Este curso es ahora un repositorio de problemas y soluciones que se planea comenzar a investigar usando técnicas de análisis de código fuente.

Por ahora se puede afirmar que el uso de pruebas automáticas para evaluar cursos de Programación de Computadores ha resultado una estrategia efectiva en el contexto de Ingeniería de Sistemas de la UFPS, que ayuda a reducir la tasa de reprobación y a mejorar las calificaciones, lo cual es consecuencia de la motivación que genera en los estudiantes el abordar los problemas como un reto de aprendizaje y no como una obligación, lo que se corresponde con el enfoque educativo de aprender haciendo y de aprendizaje basado en retos.

REFERENCIAS

- [1] M. Andreessen, “Why Software Is Eating The World”, *The Wall Street Journal*, 20 agosto 2011.
- [2] Ministerio de Tecnologías de la Información y las Comunicaciones de Colombia (Mintic), *Estudio Exploratorio Prospectiva de la Industria TI en Colombia*. Mintic, 2015. Disponible en: <http://fedesoft.org/estudios/EstudiodeProspectiva2015.pdf>

- [3] Federación Colombiana de la Industria del Software y Tecnologías Informáticas Relacionadas (Fedesoft), “Informe de Caracterización del sector de Software y Tecnologías de la Información en Colombia“, 2015.
- [4] Ministerio de Tecnologías de la Información y las Comunicaciones de Colombia (Mintic), “Plan Nacional de Ciencia, Tecnología e Innovación 2017-2022“, 2016.
- [5] Mintic y Fedesoft, “Estudio de Salarios del Sector de Software y TI de Colombia“. 2015.
- [6] Y. Bosse and M. A. Gerosa, “Why is Programming so difficult to learn?: Patterns of difficulties related to Programming learning mid-stage”, *SIGSOFT Softw Eng Notes*, vol. 41, No. 6, pp. 1-6, 2016.
- [7] R. Pea and D.-M. Kurland, *On the Cognitive Prerequisites of Learning Computer Programming*. Washington, DC: National Inst. of Education (ED), 1983.
- [8] G. M. Weinberg, *The Psychology of Computer Programming*. New York: John Wiley & Sons, Inc., 1985.
- [9] M. Kölling, “The problem of teaching object-oriented programming Part I: Environments”, *J. Object-Oriented Program.*, vol. 11, No. 9, pp. 6-12, 1999.
- [10] J. Villalobos y R. Casallas, “Looking for a new approach to teach/learn a first computer-programming course”, *International Conference on Engineering and Computer Education (ICECE)*, Madrid, España, 2005.
- [11] J. Villalobos y R. Casallas, “Proyecto CUP12: Un enfoque multidimensional frente al problema de enseñar y aprender a programar”, *Rev. Investig. UNAD*, vol. 8, N.º 2, 2009.
- [12] R. Queirós y J. P. Leal, “Programming Exercises Evaluation Systems: An Interoperability Survey”, in *Proceedings of the 4th International Conference on Computer Supported Education (CSEDU-2012)*, pp. 83-90.
- [13] F. B. Flórez, R. Casallas, M. Hernández, A. Reyes, S. Restrepo and G. Danies, “Changing a generation’s way of thinking: Teaching computational thinking through programming”, *Rev. Educ. Res.*, vol. 87, No. 4, pp. 834-860, 2017.
- [14] G. Y. Lázaro, A. M. Delgado y F. H. Vera, “Desarrollo e implementación de un marco de trabajo para el entrenamiento en Programación Competitiva”, *Rev. Univ. Cienc. Tecnol.*, vol. 20, N.º 79, pp. 69-74, 2016.
- [15] A. Chan, J. Cao, C.-K. Liu and W. Cao, “Design and Implementation of VPL: A Virtual Programming Laboratory for Online Distance Learning”, in *Advances in Web-Based Learning - ICWL 2003*, vol. 2783, pp. 509-519.
- [16] C. Douce, D. Livingstone and J. Orwell, “Automatic Test-based Assessment of Programming: A review”, *J. Educ. Resour. Comput.*, vol. 5, No. 3, sep. 2005.
- [17] P. Ihanola, T. Ahoniemi, V. Karavirta and O. Seppälä, “Review of recent systems for Automatic Assessment of Programming Assignments”, in *Proceedings of the 10th Koli Calling International Conference on Computing Education Research*, New York, USA, 2010, pp. 86-93.
- [18] J. C. Rodríguez, E. R. Royo y Z. J. Hernández, “VPL: Laboratorio Virtual de Programación para Moodle”, en *Jornadas de Enseñanza Universitaria de la Informática*, Santiago de Compostela, 2010.
- [19] Instituto Tecnológico y de Estudios Superiores de Monterrey (ITESM), *Edu Trends - Aprendizaje Basado en Retos*. Observatorio de Innovación Educativa, Tecnológico de Monterrey, 2016. [En línea]. Disponible en: <https://observatorio.tec.mx/edutrendsabr>
- [20] S. Halim and F. Halim, *Competitive Programming*, 3.^a ed., Singapore, 2013.
- [21] ICPC, *ICPC History*. 2017.
- [22] ACIS, *Consolidado Maratones de Programación Colombia*. 2017. [En línea]. Disponible en: <https://acis.org.co/portal/content/consolidado-de-maratones-de-programacion>

- [23] Red de Programación Competitiva (RPC), Sitio Red de Programación Competitiva. 2017. [En línea]. Disponible en: <http://redprogramacioncompetitiva.com/>
- [24] Colombian Collegiate Programming League (CCPL). 2017. [En línea]. Disponible en: <https://www.programmingleague.org/>
- [25] A. M. Delgado y G. Y. Lázaro, Desarrollo e implementación de un marco de trabajo para el grupo de estudio en programación competitiva del programa de ingeniería de sistemas de la Universidad Francisco de Paula Santander (archivo electrónico). San José de Cúcuta: Universidad Francisco de Paula Santander, 2018.
- [26] S. K. Robinson and I. S. Torsun, “The Automatic Measurement of the Relative Merits of Student Programs”, *SIGPLAN Not*, vol. 12, No. 4, pp. 80-93, abr. 1977.
- [27] M. J. Rees, “Automatic Assessment Aids for Pascal Programs”, *SIGPLAN Not*, vol. 17, No. 10, pp. 33-42, oct. 1982.
- [28] J. C. Rodríguez, E. R. Royo and Z. J. Hernández, “Fighting plagiarism: Metrics and methods to measure and find similarities among source code of computer programs in VPL”, in *Proc. EDULEARN11*, 2011, pp. 4339-4346.
- [29] K. Beck, *Test Driven Development: By Example*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 2002.
- [30] R. C. Martin, “Professionalism and Test-Driven Development”, *IEEE Softw*, vol. 24, No. 3, pp. 32-36, may 2007.
- [31] J. C. Rodríguez del Pino, “Sitio Web VP”, 2017. [En línea]. Disponible en: https://moodle.org/plugins/stats.php?plugin=mod_vpl
- [32] A. C. Vélez y A. Vélez, *Creatividad e inventiva: Retos del siglo XXI*. Medellín: Universidad de Antioquia, 2014.
- [33] J. Rogalski and R. Samurçay, “Acquisition of Programming Knowledge and Skills”, in *Psychology of Programming*, J. M. Hoc, T. R. G. Green, R. Samurçay and D. J. Gilmore, Eds., London: Academic Press, 1990, pp. 157-174.
- [34] R. Scherer, “Learning from the Past—The Need for Empirical Evidence on the Transfer Effects of Computer Programming Skills”, *Front. Psychol.*, vol. 7, p. 1390, 2016.
- [35] P. J. Clarke, D. Davis, T. M. King, J. Pava and E. L. Jones, “Integrating Testing into Software Engineering Courses Supported by a Collaborative Learning Environment”, *Trans. Comput. Educ.*, vol. 14, No. 3, pp. 18:1-18:33, oct. 2014.
- [36] S. A. Brian, R. N. Thomas, J. M. Hogan and C. Fidge, “Planting Bugs: A System for Testing Students’ Unit Tests”, en *Proceedings of the 2015 ACM Conference on Innovation and Technology in Computer Science Education*, New York, USA, 2015, pp. 45-50.
- [37] Y. Bosse and M. A. Gerosa, “Difficulties of Programming Learning from the point of view of students and instructors”, *IEEE Lat. Am. Trans.*, vol. 15, No. 11, pp. 2191-2199, 2017.
- [38] Y. Bosse, D. Redmiles and M. A. Gerosa, “Pedagogical Content for Professors of Introductory Programming Courses”, in *Proceedings of the 2019 ACM Conference on Innovation and Technology in Computer Science Education*, 2019 pp. 429-435.
- [39] M. Kölling, B. Quig, A. Patterson and J. Rosenberg, “The BlueJ system and its pedagogy”, *J. Comput. Sci. Educ. Spec. Issue Learn. Teach. Object Technol.*, vol. 13, No. 4, 2003.
- [40] D. Barners and M. Kölling, *Object First with Java, A practical Introduction using BlueJ*. 6th ed., Pearson, 2016.

Capítulo IV

Criterios y Métricas para Evaluar la Seguridad en Aplicaciones Web: Metodología MESW

Alexánder Barinas López, abarinas@jdc.edu.co

Docente-investigador, Fundación Universitaria Juan de Castellanos, Tunja-Colombia

John A. Bohada, jbohada@jdc.edu.co

Docente-investigador, Fundación Universitaria Juan de Castellanos, Tunja-Colombia

Iván Andrés Delgado, idelgado@jdc.edu.co

Docente-investigador, Fundación Universitaria Juan de Castellanos, Tunja-Colombia

I. INTRODUCCIÓN

En la actualidad, la información es el activo más importante para el desarrollo de las actividades en una organización. Con la masificación de internet las aplicaciones web son cada vez más populares y complejas, y brindan abundantes beneficios a las empresas en cuanto a posicionamiento en el mercado, publicidad, e información sensible para llevar a cabo la ejecución de sus procesos.

Además, es innegable que cada día las organizaciones dependen en mayor medida de la información y la tecnología, y que ambas, actualmente, soportan los sistemas de información. Pero no siempre ha sido así. En años anteriores, la protección de la información era más sencilla, ya que las grandes plataformas de datos actuaban de forma independiente sin conectividad alguna, las arquitecturas existentes eran totalmente centralizadas y las terminales tenían capacidades de procesamiento limitadas; en cambio, en la actualidad, esa posibilidad se ha incrementado notablemente con la diversidad de redes y plataformas tanto internas como externas, y la información está cada vez más expuesta a las diferentes amenazas y ataques informáticos [1] [2], cualquier dispositivo conectado a la internet corre el riesgo de ser atacado desde computadoras, servidores, teléfonos celulares, entre otros; asimismo, la mayoría de las aplicaciones y servicios web son susceptibles a un conjunto de ataques debido a las vulnerabilidades originadas en los defectos del diseño y la implementación de las aplicaciones, en la programación descuidada de las

rutinas, en la escasa implementación de medidas de control de acceso o en la falta de validación de los datos de entrada o incluso en ataques originados por ingeniería social [3].

Por ello, toda organización debe contar con políticas e instrumentos que garanticen la valoración de los riesgos a los cuales está expuesta su información, implementar procedimientos de solución que minimicen la afectación que estos riesgos puedan ocasionar, y crear o fortalecer las áreas de TI que permitan hacer seguimiento y control para garantizar la integridad, confidencialidad y protección de su activo más valioso: la información [4].

Por tanto, el objetivo de esta investigación es presentar el diseño de una metodología de evaluación con base en criterios y métricas de acuerdo con normas y estándares de seguridad, que permita estimar el nivel de seguridad de una aplicación web, de manera que les sirva a los desarrolladores para tomar decisiones y medidas preventivas en el desarrollo de sus aplicaciones.

II. MATERIALES Y MÉTODOS

Para el desarrollo de esta investigación, en primer lugar, se realizó una indagación bibliográfica que permitiese una contextualización de este artículo orientada a la seguridad web a partir de su descripción y de los proyectos, aplicaciones y trabajos desarrollados. En segundo lugar, y basado en dicha contextualización, se propone una novedosa metodología que permite evaluar o

estimar el nivel de seguridad de una determinada aplicación web. Dicha metodología se describe y, además, muestra el resultado de su ejecución en la aplicación, fase por fase.

III. RESULTADOS

A. Fundamento Teórico

La seguridad es un aspecto primordial que todo sistema debe incorporar con el fin de proteger los activos de cualquier amenaza. Es importante conocer los rasgos inherentes a la seguridad de la información y la protección de datos. Sin embargo, antes de abordar la problemática existente en las aplicaciones web, es necesario describir los fundamentos teóricos más relevantes que servirán como base para el desarrollo de la metodología propuesta, los cuales se presentan a continuación:

1. *Seguridad Web*: Actualmente, internet está integrado por un conjunto de servicios y aplicaciones que funcionan en la red, desde comercio electrónico, transferencia de archivos, transacciones que mueven enormes cantidades de dinero, búsquedas en la web, correo electrónico, hasta sitios de redes sociales que contienen información importante y confidencial de sus miembros [5].

Mientras aumenta de forma considerable el número de conexiones a la red, así como la migración de aplicaciones y de usuarios a los servicios de internet, crecen las amenazas y sus consecuencias y las medidas tomadas para contrarrestarlas y por tanto la demanda de servicios seguros en la web (ver Tabla 1). Sin lugar a dudas, el aspecto más crítico de la seguridad en internet está en los actores que intervienen directamente en su uso y en la prestación de los servicios, es decir, los usuarios y los servidores web.

Tabla 1. Amenazas, consecuencias y contramedidas a la seguridad web

	Amenazas	Consecuencias	Contramedidas
Integridad	<ul style="list-style-type: none"> - Virus, puertas traseras y caballos de Troya - Modificación de datos de usuario - Codificación descuidada - Reemplazos o borrados accidentales - Modificación de tráfico de mensajes en tránsito 	<ul style="list-style-type: none"> - Pérdida de información - Compromiso de la máquina - Vulnerabilidad ante todas las otras amenazas 	<ul style="list-style-type: none"> - Cifrado de datos - Sistemas de detección de intrusos - Verificaciones de resúmenes
Confidencialidad	<ul style="list-style-type: none"> - Monitoreo de la red - Captura de información acerca de la configuración de la red - Acceso no autorizado a archivos - Acceso remoto no autorizado a bases de datos 	<ul style="list-style-type: none"> - Pérdida de información - Pérdida de confidencialidad 	<ul style="list-style-type: none"> - Cifrado de datos - Clasificación de datos - Proxies Web
Disponibilidad	<ul style="list-style-type: none"> - Destrucción de procesos de usuario - Ataques de denegación de servicio (Denial Of Service o DoS) - Inhabilitar disco o memoria - Cortes en las líneas de comunicaciones 	<ul style="list-style-type: none"> - Disrupción - Anonimidad - Evita que los usuarios hagan su trabajo 	<ul style="list-style-type: none"> - Implantación de sólidos mecanismos de control de acceso - Implantación de sistemas redundantes en aplicaciones críticas - Copias de respaldo de la información vital
Autenticación	<ul style="list-style-type: none"> - Suplantación de usuarios legítimos - Falsificación de datos 	<ul style="list-style-type: none"> - Representación errónea del usuario - Creencia en la validez de la información falsa 	<ul style="list-style-type: none"> - Técnicas criptográficas - Autenticación de usuarios - Autorización de usuarios
No repudio	<ul style="list-style-type: none"> - Posibilidad de repudio - Falsificación del origen del envío 	<ul style="list-style-type: none"> - Robos o sabotaje - Fraude electrónico 	<ul style="list-style-type: none"> - Cifrado - Firma digital o certificado digital - Control de acceso

Fuente: Tomado de [6].

Como es sabido, las fallas de seguridad en los servidores web en muchos casos son ocasionadas por descuidos en los mecanismos de protección, fallas en los accesos de control y protección a las aplicaciones y a los datos [7], sin embargo, la razón de la mayoría de fallas de seguridad son las aplicaciones que no se desarrollan de manera profesional, y en varios casos sus autores no tienen conocimiento profundo de los problemas asociados con la seguridad.

Asimismo, los tres propósitos que con mayor frecuencia persiguen los atacantes a sitios web son los siguientes [8], [9]:

- Dejar el sitio fuera de servicio, lo que se conoce como Denegación de Servicio (DoS, por sus siglas en inglés). Los motivos que llevan a este tipo de ataque pueden ser: dejar sin acceso a los usuarios legítimos e imposibilitar el acceso a los servicios y recursos de una organización, o simplemente para atentar contra la disponibilidad del sistema.
- Atentar contra la integridad del sistema modificando o alterando sin autorización los programas, lo que conduce a provocar fallos en el sistema que afectan los procesos y la imagen de una organización.
- Obtener información no autorizada con el objetivo de captar información sensible de las personas, y realizar espionaje de interés para la competencia, como archivos de cuentas bancarias, lo que afecta la confidencialidad de la información.

Teniendo en cuenta que las causas de un ataque son diversas, desde un pirata informático que busca poner a prueba sus destrezas técnicas y lograr captar información sensible, hasta los actos de ciberterroristas que pueden llegar a impedir el normal funcionamiento de una organización. De esta manera, el éxito de un ataque va a depender del minucioso análisis al comportamiento del sitio web, de las debilidades del sitio o de un pobre control de acceso.

2. *Iniciativas para la seguridad web:* Actualmente hay varias iniciativas que velan por prevenir y ayudar a la seguridad web de las organizaciones [4], ejemplo de ello son OWASP, NIST 800-115, ISSAF y MAGERIT. El proyecto Abierto de Seguridad de Aplicaciones Web (OWASP, por sus siglas en inglés) [10], [11], es uno de los que más relevancia tienen y es una comunidad de código abierto dedicada a determinar y combatir las causas que hacen que el software sea inseguro. Este organismo sin ánimo de lucro conforma una serie de guías y proyectos relacionados con la implementación de la seguridad principalmente en entornos web, además concentra sus esfuerzos en la gestión de metodologías de apoyo a las organizaciones para que desarrollen, adquieran y mantengan aplicaciones en las que se pueda confiar. Para el caso de las aplicaciones web, la iniciativa OWASP presenta una guía importante a tener en cuenta para la seguridad de este tipo de aplicaciones, ya que este proyecto continuamente realiza estudios de las vulnerabilidades y ataques. En la Tabla 2 se describen las principales vulnerabilidades publicadas por este proyecto [10] [11].

Tabla 2. Principales vulnerabilidades en aplicaciones web

Vulnerabilidad	Descripción
A1 - Inyección	Corresponde a las inyecciones de código, siendo las inyecciones SQL unas de las más comunes.
A2 – Pérdida de autenticación	Corresponde al mal manejo de las sesiones en aquellas aplicaciones que utilizan autenticación.
A3 – Exposición de datos sensibles	Se refiere a la protección incorrecta de datos críticos, por ejemplo, números de tarjetas de crédito, contraseñas, etc.
A4 – XML External Entities (XXE)	Sucede cuando las entidades externas se pueden utilizar para divulgar archivos internos, posibilitando la ejecución remota de código y ataques de denegación de servicio.

Vulnerabilidad	Descripción
A5 – Pérdida de control de acceso	Puede derivar en un acceso no autorizado a información crítica debido a fallos en la restricción a los usuarios autenticados.
A6 – Configuración de seguridad incorrecta	Corresponde a configuraciones no adecuadas que pueden impactar en la seguridad de la propia aplicación.
A7 – Secuencia de comandos en sitios cruzados (XSS)	Ocurre cuando existe validación pobre de la información ingresada por el atacante.
A8 – Deserialización insegura	Esta vulnerabilidad puede permitir la ejecución remota de código en servicios web, como ataques de repetición, de inyección y de escalamiento de privilegios.
A9 – Uso de componentes con vulnerabilidades conocidas	Corresponde a la explotación de librerías, <i>frameworks</i> y otros componentes vulnerables por parte de un atacante, con el fin de obtener acceso o combinar con otros ataques.
A10 – Registro y monitorización insuficientes	Permite a los atacantes mantener la persistencia, pivotar hacia más sistemas y manipular, extraer o destruir datos.

Fuente: Tomado de [10].

Por su parte, la guía técnica para la evaluación y pruebas de seguridad de la información (NIST SP 800-115, por sus siglas en inglés), presenta cuatro fases: planificación, descubrimiento, ejecución y reportes. En la fase de planificación se realiza una aprobación de la prueba por parte de la gerencia y se identifican su alcance y objetivos; en la fase de descubrimiento se recopila la información y se analizan las vulnerabilidades que se presentaron; en la fase de ejecución del ataque se comprueban las vulnerabilidades previamente detectadas y, si es necesario, se hace un proceso de mitigación de las mismas; finalmente, en la fase de reportes o presentación del informe se documentan todas las fases y las acciones o resultados que se van obteniendo de cada una de ellas y se hacen las posibles recomendaciones. Cada una de las fases

se hace en una secuencia, al llegar a la segunda y tercera fases (ataque y descubrimiento) se presenta un bucle de retroalimentación, momento en el cual se deben realizar las pruebas y el análisis sobre múltiples sistemas para determinar el nivel de acceso que puede tener un atacante.

De igual forma, la plataforma para la evaluación de la seguridad de los sistemas de información (ISSAF, por sus siglas en inglés), se compone de tres fases y nueve pasos cíclicos para el desarrollo de su proceso, las fases corresponden a: planificación y preparación, evaluación, y reportes y limpieza y destrucción de información. En la fase de planificación y preparación se deben planear los aspectos más importantes de las pruebas que se llevarán a cabo antes de ejecutarse y firmar un acuerdo; en la fase de evaluación se realiza un enfoque por capas en el que cada una de ellas presenta un mayor nivel de acceso a los activos de información, y en la fase final se debe generar un informe con los resultados obtenidos y las recomendaciones pertinentes. Adicionalmente, la información que se almacenó en los sistemas de prueba debe ser eliminada o, en su defecto, mencionada en el informe técnico para su posterior eliminación.

Finalmente, la Metodología de Análisis y Gestión de Riesgos de los Sistemas de Información (Magerit) fue desarrollada por un ente gubernamental como respuesta a la percepción de que la administración, y en general toda la sociedad, depende de forma creciente de las tecnologías de la información para el cumplimiento de su misión [12]; además, se afirma que el análisis del riesgo no es más que un proceso sistemático para estimar la magnitud de los riesgos a los que está expuesta una organización, y este análisis permite determinar cómo es, cuánto vale y cuán protegidos se encuentran los activos informáticos. Los objetivos que Magerit persigue son: concienciar a los gerentes de las organizaciones sobre los riesgos que tiene la información y la importancia de gestionarlos, ofrecer un método sistemático para analizar los riesgos derivados del uso de tecnologías de la información y la comunicación, descubrir los riesgos y mantenerlos bajo control, y preparar la organización para

procesos de evaluación, auditoría, certificación o acreditación, según sea el caso.

B. Referentes en Seguridad Web

Para llevar a cabo esta investigación, se referenciaron algunos documentos y artículos relacionados con el tema de seguridad, que corresponden a estudios adelantados en torno a la problemática y que permitieron dar soporte a la presente investigación.

En [13] se describen los pasos a seguir para establecer un modelo SGSI en las instituciones de educación superior, teniendo como base el estándar ISO 27001.

En [14] los autores proponen el diseño y la construcción de un modelo para evaluar la seguridad en ambientes virtuales de aprendizaje (AVA), a partir de la identificación de criterios de seguridad propuestos en guías, normas y estándares. Una vez establecidos los criterios, se analizaron métricas que permiten cuantificar dicha relación para facilitar la validación del modelo propuesto, el cual fue aplicado a los ambientes Moodle y Dokeos, dos plataformas de código abierto y distribución libre, lo que permite llevar a cabo un juicio de valor confiable tanto cuantitativo como cualitativo sobre el estado de seguridad de los AVA [14].

En [15] se analizan los principales ataques que se han producido en los últimos años y que han comprometido las aplicaciones web, sus bases de datos, los usuarios, entre otros; asimismo, los autores investigan la tendencia que tienen los ataques a las aplicaciones web en las diferentes categorías, con el objetivo de saber qué vulnerabilidades están comúnmente en riesgo y utilizar ese análisis en la creación de aplicaciones web más seguras en el futuro y poder guiar a los desarrolladores web a tomar medidas preventivas [15].

En [16] se presentan estudios sobre la detección de vulnerabilidades Cross Site Scripting (XSS), pero muy pocos se han centrado en su eliminación, por lo cual en este documento los autores enfatizan

tanto en los métodos como en las herramientas para eliminar vulnerabilidades XSS; los autores dentro del estudio también muestran una descripción de las diferentes técnicas o modelos capaces de disminuir o filtrar en el momento vulnerabilidades en aplicaciones web, e invitan a los investigadores a centrarse más en su identificación y eliminación antes de hacer una implementación en las web [16].

C. Metodología para Evaluar la Seguridad Web (MESW)

Antes de describir la metodología y cada una de sus fases, se indica qué aplicación se tomó como base de prueba para generar los resultados presentados en esta sección.

1. Aplicación web para realizar evaluación de seguridad: Para realizar las pruebas de seguridad, se tomó como referente una aplicación web desarrollada en el grupo de investigación Ciencia, Innovación y Tecnología (CIyT), de la Facultad de Ingeniería y Ciencias Básicas de la Fundación Universitaria Juan de Castellanos. Por seguridad, el dominio de dicha aplicación no será referenciado.

La aplicación trata de una solución web para una tienda de venta de artesanías en línea. La aplicación cuenta con dos tipos de usuario: usuario general y usuario administrador del sitio. El usuario general podrá registrarse e iniciar sesión en el sitio, consultar los artículos más recientes, los artículos por su tipo, artículos por precio (menor a mayor), buscar artículos, ver el detalle de cada artículo, agregar un artículo al carrito de compras y realizar la compra. Por otro lado, el usuario administrador cuenta con privilegios tales como: ingresar al sitio, consultar los artículos más recientes, artículos por su tipo, artículos por precio (menor a mayor), buscar artículos, ver el detalle de cada artículo, tener un menú de administración, registrar un nuevo artículo, ver el listado de artículos, de clientes y de ventas, entre otras funcionalidades.

Desde el punto de vista técnico, la aplicación fue desarrollada bajo un servidor ApacheServer, lenguaje web PHP5, base de datos MySQL Server.

2. Descripción y uso de la metodología MESW:

La metodología para evaluar la seguridad de aplicativos web (MESW) está compuesta por cinco fases: estado del sistema, diseño de pruebas, identificación de vulnerabilidades, evaluación de la seguridad y análisis de resultados (Figura 1).

Las fases de MESW son consecutivas y están compuestas por diferentes actividades que se llevan a cabo para medir o estimar el nivel de seguridad de una determinada aplicación web, de manera que sirva a los desarrolladores para tomar decisiones y medidas preventivas en el desarrollo de sus aplicaciones.

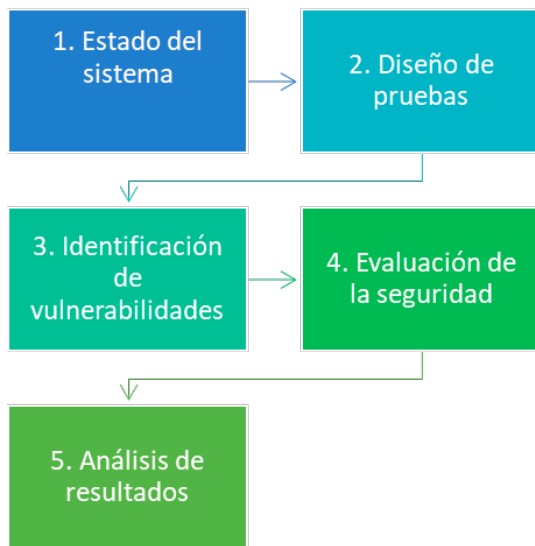


Fig. 1. Diseño metodológico de evaluación de seguridad

Fase 1. Estado del sistema: En esta fase se realiza la identificación de la aplicación que se va a evaluar, asimismo, se lleva a cabo el reconocimiento y registro de las características del equipo donde se encuentra alojada la aplicación web, es decir, el servidor web. Las características del servidor web donde se encuentra alojada la aplicación objeto de evaluación son las que se muestran en la Tabla 3.

Tabla 3. Características del servidor web

Hardware/ Software	Descripción
CPU	Intel Core i5 1.7 GHz - 2.4 GHz
RAM	8 Gb
Sistema operativo	Windows 10 Home
Stack	WAMP
Servidor web	Apache Server
Base de datos	MySQL Server
Lenguaje web	PHP 5
Aplicación web	

Fuente: Elaboración propia.

Posteriormente, se llevan a cabo las pruebas de seguridad, las cuales se hacen desde un equipo local con sistema operativo Windows y ejecutando las herramientas de seguridad propuestas y descritas en la siguiente fase.

Fase 2. Diseño de pruebas: En esta fase se deben realizar tres actividades: 1: Identificar herramientas de seguridad que permitan evaluar, inicialmente, la seguridad de la aplicación web. 2: Diseño de métricas y criterios de seguridad de acuerdo con normas y estándares de seguridad. 3: Con base en las métricas propuestas, se establecen valores de medición según las normas o estándares usados en la actividad 2.

En este sentido, el desarrollo de las tres actividades correspondió a lo siguiente:

Actividad 1: Se realizó una indagación de herramientas que permitiesen evaluar la seguridad web del aplicativo objeto de estudio. En este sentido, se identificaron las siguientes tres herramientas: SQLMap (herramienta automática de inyección SQL) [17], W3af (*frame* de testeo de auditoría y ataques a servicios web) [18], IronWASP (plataforma avanzada de testeo de seguridad web) [19].

Actividad 2: Para el desarrollo de las métricas y criterios de seguridad, se tomó como base la lista de vulnerabilidades definidas en el proyecto OWASP descritas en la Tabla 2 y los tres criterios de seguridad definidos en la norma ISO 27001 [20]: confidencialidad (C), integridad (I) y disponibilidad (D).

La Tabla 4 describe las métricas utilizadas, las ponderaciones definidas por cada métrica y los valores que se tuvieron en cuenta, según OWASP, para evaluar el nivel de seguridad de la aplicación web.

Tabla 4. Descripción de ponderaciones

Métrica	Ponderación	Valores teniendo en cuenta OWASP
Gestión de configuraciones	15 %	0 – 15
Gestión de codificación y validación de entrada	24 %	0 – 24
Gestión de control de acceso y autenticación	16 %	0 – 16
Gestión de sesiones y usuario	16 %	0 – 16
Gestión de errores y excepciones	6 %	0 – 6
Gestión de datos sensibles	23 %	0 – 23

Fuente: Elaboración propia.

Actividad 3: Por cada métrica se debe hacer la evaluación correspondiente, donde se deben identificar los aspectos para evaluar y por cada aspecto, se establecen valores según el impacto de la vulnerabilidad de acuerdo con los tres criterios mencionados anteriormente (I, C, D). Ejemplo de ello es la evaluación que se describe en la Tabla 5, donde se tomó como base la métrica de gestión de codificación y validación de entrada.

Tabla 5. Métrica de gestión de codificación y validación de entrada

Nº	Aspectos a evaluar	I.	C.	D.	SÍ/NO
4	¿Existen vulnerabilidades de inyección de SQL?	9	9	8	
5	¿Existen Vulnerabilidades Remote File Inclusion?	7	7	7	
6	¿Existen vulnerabilidades Local File Inclusion?	5	5	6	
7	¿Existen vulnerabilidades Cross Site Scripting?	3	3	3	
Total		24	24	24	

Fuente: Elaboración propia.

De acuerdo con la métrica descrita en la Tabla 5, sus cuatro aspectos se tomaron para el proceso de evaluación, y los valores fueron calculados para cada criterio (I, C, D) de acuerdo con el impacto de la vulnerabilidad en la red en relación con los aspectos de explotabilidad, prevalencia y detección de la vulnerabilidad según OWASP. Para nuestro caso, se tomó una escala de 10 a 1, donde 10 es el mayor impacto de la vulnerabilidad y 1 el menor impacto. Para dicha métrica se obtuvo un valor de 24 sobre el total.

Fase 3. Identificación de vulnerabilidades:

En esta fase, para identificar las vulnerabilidades de la aplicación web objeto de estudio, se deben ejecutar las herramientas de análisis de seguridad mencionadas en la fase de pruebas, actividad 1: SQLMap, W3af e IronWASP, desde un equipo local, analizando la aplicación en el servidor web que es el objetivo de análisis de prueba.

En esta etapa se aplica el diseño de las pruebas definido en la fase anterior, para posteriormente realizar el análisis de las vulnerabilidades detectadas, tanto con las pruebas automáticas como con las manuales. Ejemplo del proceso de ejecución

es el que se muestra en la Figura 2, donde se está utilizando la herramienta W3af para el análisis de las vulnerabilidades de la aplicación objeto de estudio. Para este caso, la herramienta no detectó ninguna de las vulnerabilidades descritas anteriormente en la Tabla 2.

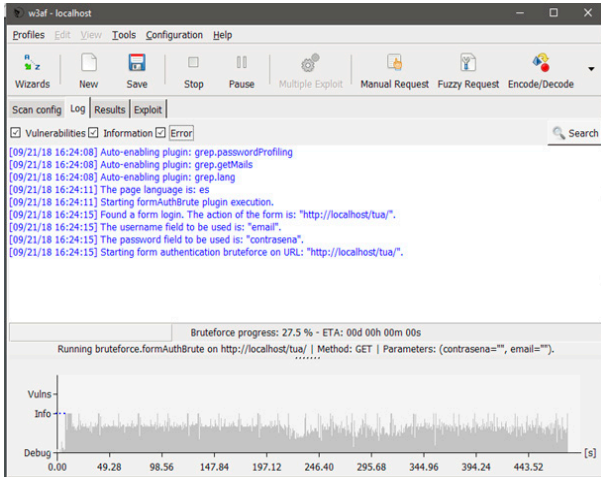


Fig. 2. Ejecución herramienta W3af

Fase 4. Evaluación de seguridad: En esta fase, para hacer seguimiento es necesario diseñar una plantilla (Plantilla_Métrica de evaluación de seguridad), que permita al evaluador determinar si existe o no la vulnerabilidad según los aspectos definidos en la fase anterior. La Figura 3 muestra la plantilla que se desarrolló para esta investigación.

En este caso, la plantilla cuenta con 20 aspectos, los cuales están representados en seis métricas, cada una de las cuales tiene un valor asignado de acuerdo con el impacto de la vulnerabilidad según los criterios de Integridad, Disponibilidad y Confidencialidad definidos por la norma ISO 27001 para la Gestión de Seguridad de la Información. Los valores se calculan automáticamente cuando el agente evaluador determina si existe o no una vulnerabilidad (con SÍ o NO) y se registra la respectiva observación.

N	Métricas	SI/NO/NA	Observación
GESTIÓN DE CONFIGURACIONES			
1.	¿Existe algún software sin actualizar? Incluye el Sistema Operativo, Servidor Web/Aplicación, DBMS, aplicaciones, y librerías del código.	NO	
2.	¿Tiene habilitada o instalada alguna característica innecesaria? (por ejemplo puertos, servicios, paginas, cuentas, privilegios).	NO	
3.	¿Están las cuentas por defecto y las contraseñas aun habilitadas y sin cambiar?	NO	
GESTIÓN DE CODIFICACIÓN Y VALIDACIÓN DE ENTRADA			
4.	¿Existen vulnerabilidades de inyección de SQL?	NO	
5.	¿Existen vulnerabilidades Remote File Inclusion (RFI)?	NO	
6.	¿Existen vulnerabilidades Local File Inclusion (LFI)?	NO	
7.	¿Existen vulnerabilidades Cross Site Scripting (XSS)?	NO	
GESTIÓN DE CONTROLES DE ACCESO Y AUTENTICACIÓN			
8.	¿Existen vulnerabilidades CSRF (Cross Site Request Forgery)?	NO	
9.	¿Carece de restricciones de tiempo, número o Captcha para las peticiones de autenticación de los usuarios de la plataforma?	SI	
10.	¿La aplicación permite el listado de ficheros internos o confidenciales como archivos de configuración, backup, temporales, etc.?	NO	
11.	¿La aplicación gestiona de manera inadecuada el nivel de acceso a recursos?	NO	
GESTIÓN DE SESIONES Y DE USUARIOS			
12.	¿La aplicación establece un tiempo de validez para la sesión del usuario?	SI	
13.	¿Se gestiona de manera adecuada el identificador de mi sesión?	SI	
14.	¿El identificador de la sesión es almacenado y transportado en otro lugar que no sea la URL?	SI	
GESTIÓN DE ERRORES Y EXCEPCIONES			
15.	¿El Sistema sigue en funcionamiento aun después de alguna excepción o error de su dominio evitando así una denegación de servicio?	SI	
16.	¿El manejo de errores no deja en evidencia rastros de las capas de aplicación u otros mensajes de error demasiado informativos a los usuarios como rutas absolutas, configuraciones, etc.?	SI	
GESTIÓN DE DATOS SENSIBLES			
17.	¿La aplicación guarda información sensible en el navegador del usuario (local storage, cookies, etc.)?	NO	
18.	¿Los datos sensibles se almacenan en texto claro a largo plazo, incluyendo sus respaldos?	NO	
19.	¿Los datos sensibles se transmiten en texto claro, interna o externamente?	NO	
20.	¿Se utiliza algún algoritmo criptográfico débil o antiguo?	NO	

Fig. 3. Plantilla de evaluación de seguridad

Fase 5. Resultados: La fase de resultados permite obtener los valores de cada una de las métricas. En esta fase se presentan y analizan los resultados obtenidos de la plantilla de métricas, lo que permite conocer el nivel de seguridad de la aplicación.

Por tanto, una vez se hayan evaluado todas las métricas con sus respectivos aspectos se procede a establecer y obtener un resultado de la medida o el nivel de seguridad por cada uno de sus criterios: integridad, confidencialidad y disponibilidad. La Tabla 6 y la Figura 4 presentan los resultados que se obtuvieron una vez aplicada la plantilla de métricas definida.

El resultado de la evaluación de la aplicación web se obtiene del promedio de los tres criterios y luego se determina el nivel de seguridad de forma cualitativa, según los rangos definidos en la Tabla 7.

Tabla 6. Resultados de la evaluación a la aplicación web

Métrica	Resultado
Integridad	96
Confidencialidad	96
Disponibilidad	88
Seguridad aplicación web	93,33
Nivel	Alto

Fuente: Elaboración propia.

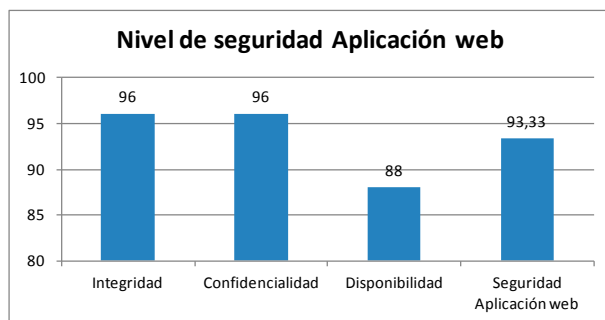


Fig. 4. Resultado gráfico de la evaluación a la aplicación web

De acuerdo con la plantilla de evaluación, el resultado fue de 93,33, y según los rangos establecidos para medir cualitativamente (los rangos son propios de la metodología MESW), tal y como muestra la Tabla 7, el resultado final fue Nivel de seguridad ALTO.

Tabla 7. Valores para cada escala

Valor cualitativo	Valor cuantitativo
Nivel bajo	0 – 60
Nivel medio	61 – 90
Nivel alto	91 – 100

Fuente: Elaboración propia.

Esto significa que automáticamente en la plantilla se calculan los datos teniendo en cuenta fórmulas y funciones para obtener una valoración final con respecto al estado de seguridad de la aplicación; en este caso, como se mencionó anteriormente, la aplicación web se encuentra en un nivel de seguridad alto, es decir, es un entorno confiable y mantiene mecanismos de seguridad robustos, sin embargo, carece de restricciones de tiempo, número o *captcha* para las peticiones de autenticación de los usuarios (numeral 9 de la plantilla de evaluación de la seguridad).

3. *Otras pruebas de uso de MESW:* De igual forma, y para determinar el comportamiento de la metodología MESW, ésta también se utilizó para evaluar ambientes virtuales de aprendizaje (AVA), se aplicó a Moodle y Dokeos, dos plataformas de código abierto y distribución libre, y permitió llevar a cabo un juicio de valor confiable tanto cuantitativo como cualitativo sobre el estado de seguridad de los AVA.

Al comparar los resultados de evaluación de seguridad a los AVA Moodle y Dokeos con el modelo propuesto, se pudo establecer un valor cuantitativo y cualitativo del nivel de seguridad de los AVA, fue así que Moodle obtuvo una calificación de 91,33 puntos para una escala de nivel ALTO, mientras que Dokeos obtuvo un valor de 89,33 puntos y su escala fue de nivel MEDIO.

Al hacer la comparación de seguridad entre Moodle y Dokeos, se observó que ambas plataformas toman medidas en la gestión y codificación de entradas, que es un pilar de vital importancia para mantener el nivel de seguridad de las plataformas, y aunque el aplicativo Dokeos obtuvo un valor inferior a Moodle en la métrica de Gestión de controles de acceso y autenticación, mantiene un nivel de seguridad confiable para su implementación en entornos de producción.

IV. DISCUSIÓN

La metodología MESW introducida en este artículo es útil para medir el estado de seguridad de una aplicación web de una manera modular, porque permite saber cuáles son los eslabones débiles del aplicativo (configuraciones, codificación y validación de entradas, etc.), de esa manera, se puede analizar el costo-beneficio al ejecutar medidas que mitiguen posibles riesgos de seguridad.

V. CONCLUSIONES

La seguridad de la información es un problema que afecta continuamente a las organizaciones, causando pérdidas tanto económicas como en los procesos de negocio. Por tal motivo, fue necesario diseñar y construir una metodología de evaluación de seguridad teniendo en cuenta normas, estándares y organizaciones encargadas de velar por la seguridad de las aplicaciones web, las cuales fueron los insumos para su diseño. De igual forma, el diseño de un modelo de evaluación de seguridad (plantilla de la Figura 3) incrementa el carácter novedoso de la metodología MESW, dado que no existe, o al menos no se evidenció en el proceso de desarrollo de este artículo, una estructura que incorpore 6 métricas y 20 aspectos para la evaluación de la seguridad de aplicaciones web.

La metodología propuesta será de gran utilidad para prevenir posibles ataques y vulnerabilidades que se puedan presentar en una aplicación web, y a través de su uso va a permitir generar resultados que ayudarán a los desarrolladores a tomar decisiones oportunas y de esta manera poder mejorar o implementar mecanismos de seguridad y de control en el sistema.

Los resultados de la investigación muestran la importancia de contar con el diseño de una metodología de evaluación de seguridad, ya que ésta permitió evaluar la seguridad de diversas aplicaciones

web, sin embargo, se hace necesario seguir aplicando la metodología periódicamente para su evaluación, con el fin de tomar medidas correctivas contra las posibles amenazas que se presenten.

Aunque existen diversas guías, normas y estándares de seguridad concerniente a la seguridad informática y de la información a nivel general en una organización, es poco lo que existe respecto a normas de seguridad en aplicaciones web. Gracias al proyecto OWASP, el cual ha dedicado su estudio y preocupación por las vulnerabilidades más relevantes descritas en el ranquin de las 10 principales, tanto arquitectos, como desarrolladores y administradores de software y de redes tienen mayores expectativas para mejorar y aplicar buenas prácticas en la seguridad de los activos.

REFERENCIAS

- [1] V. Mouli and K. P. Jevitha, "Web Services Attacks and Security- A Systematic Literature Review", *Procedia Comp. Sci.*, vol. 93, pp. 870-877, 2016.
- [2] M. Seyyar, F. Çatak and E. Gül, "Detection of attack-targeted scans from the Apache HTTP Server access logs", *Appl. Comput. Inf.*, vol. 14, N.º 1, pp. 28-36, 2018.
- [3] S. A. Pineda, John A. Bohada y M. L. Pineda, "Ingeniería Social en Instituciones de Educación Superior", *Revista Colombiana de Tecnologías de Avanzada*, vol. 2, N.º 32, pp. 52-61, 2018.
- [4] G. V. Ríos, John A. Bohada e I. A. Delgado, "Gestión de seguridad de la información en las organizaciones", en *Investigación e Innovación en Ingeniería del Software*. Medellín: Tecnológico de Antioquia, vol. 2, 2018, pp. 111-121.
- [5] D. Aranda, J. Sánchez-Navarro and L. Mohammadi, "An overview of use, knowledge and perceptions of the Internet in Spain", *Data in Brief*, vol. 19, pp. 1498-1503, August 2018.

- [6] E. Daltabuit, E. L. Hernández, G. Mallén y J. Vázquez, *La seguridad de la información*. México: Limusa, 2007.
- [7] R. Varsha and K. P. Jevitha, “Web Services Attacks and Security. A Systematic Literature Review”, *Procedia Comp. Sci.*, vol. 93, N.º 016, pp. 870-877, 2016.
- [8] A. Prakash, M. Satish, T. Sri Sai Bhargav and N. Bhalaji, “Detection and Mitigation of Denial of Service Attacks Using Stratified Architecture”, *Procedia Comp. Sci.*, vol. 87, pp. 275-280, 2016.
- [9] A. Carlin, M. Hammoudeh and O. Aldabbas, “Defense for Distributed Denial of Service Attacks in Cloud Computing”, *Procedia Comp. Sci.*, vol. 73, pp. 490-497, 2015.
- [10] OWASP, “OWASP Top 10 – 2017. The Ten Most Critical Web Application Security Risks”, 2017 [Online]. Available: https://www.owasp.org/images/7/72/OWASP_Top_10-2017_%28en%29.pdf.pdf [Accessed: September 2018].
- [11] A. López, OWASP Testing Guide v4.0, Guía de seguridad en aplicaciones Web, marzo 2017 [En línea]. Disponible en: <https://www.owasp.org/images/1/19/OTGv4.pdf>
- [12] Ministerio de Hacienda y Administraciones Públicas de España, Secretaría General Técnica, *MAGERIT – versión 3.0. Metodología de Análisis y Gestión de Riesgos de los Sistemas de Información. Libro I – Método*. Madrid: Ministerio de Hacienda y Administraciones Públicas, octubre de 2012.
- [13] H. Alemán, “Metodología para la implementación de un SGSI en la Fundación Universitaria Juan de Castellanos, bajo la norma ISO 27001:2005”, tesis de maestría, Univ. Internacional de La Rioja, mayo 2017 [En línea]. Disponible en: http://reunir.unir.net/bitstream/handle/123456789/3129/HelenaClaraIsabel_Aleman_Novoa.pdf?sequence=1
- [14] M. Callejas, A. Alarcón and A. Barinas, “Security evaluation model for virtual learning environments”, *XI Latin American Conference on Learning Objects and Technology (LACLO)*. San Carlos, pp. 1-6, 2016. doi: 10.1109/LACLO.2016.7751773
- [15] D. Kaur and P. Kaur, “Empirical analysis of web attacks”, *Procedia Comp. Sci.*, vol. 78, pp. 298-306, 2015.
- [16] A. W. Marashdih and Z. F. Zaaba, “Cross Site Scripting: Removing Approaches in Web Application”, *Procedia Comp. Sci.*, vol. 124, pp. 647-655, 2017.
- [17] SQLmap, “Automatic SQL injection and database takeover tool”, 2018 [Online]. Available: <http://sqlmap.org/>
- [18] w3af, “Web Application Attack and Audit Framework”, 2018 [Online]. Available: <http://w3af.org/>
- [19] IronWASP, Open Source Advanced Web Security Testing Platform, [Online]. Available: <https://www.utest.com/tools/ironwasp>
- [20] Normas ISO, “ISO 27001 Gestión de Seguridad de la Información”, 2018, [En línea]. Disponible en: <https://www.normas-iso.com/iso-27001/>

Capítulo V

Caracterización y Diseño de un Modelo de Datos para la Gestión de Indicadores del Sistema de Aseguramiento de Alta Calidad de la Institución Universitaria Colegio Mayor del Cauca

Karime Alejandra Gómez - karimealejandra@unimayor.edu.co

Ingeniera Informática. Institución Universitaria Colegio Mayor del Cauca

María Camila Dueñas - mcamila@unimayor.edu.co

Ingeniera Informática. Institución Universitaria Colegio Mayor del Cauca

Fredy Alonso Vidal - fvidal@unimayor.edu.co

Decano de la Facultad de Ingeniería. Institución Universitaria Colegio Mayor del Cauca

1. INTRODUCCIÓN

Uno de los aspectos que debe caracterizar a una institución educativa, especialmente universitaria, es la forma como están organizados los procesos de auto-evaluación y autorregulación, de ahí que los sistemas de información constituyan una característica importante, que no solamente evidencia una planeación, sino que también, a través de los indicadores, permite que haya un aseguramiento de la calidad. Los sistemas de información dan cuenta de una gestión institucional y de una intención de mejora constante para lograr niveles adecuados de calidad. Hoy en día las universidades deben someterse a múltiples procesos de evaluación, adelantados por el Ministerio de Educación Nacional, como máximo organismo que supervisa la educación en Colombia, con el fin de garantizar que se cumplan los estándares de calidad.

La Institución Universitaria Colegio Mayor del Cauca, en su proceso de autoevaluación, ha venido replanteando y rediseñando procesos claves para el fortalecimiento institucional, y se ha hecho evidente la necesidad de contar con indicadores para aseguramiento de la calidad, relacionados con los diferentes procesos que se desarrollan en el campus. Por lo tanto, se pretende caracterizar esos indicadores que son la base para el aseguramiento de la calidad, identificarlos y analizarlos para realizar el modelado de una base de datos que complemente las existentes.

Este trabajo de caracterización de indicadores va a permitir un reporte de información, el cual actualmente no cumple las expectativas debido a que el crecimiento que ha tenido la Universidad en los últimos tiempos no desarrolló a la par un sistema de información que condensara la información académica, administrativa, educativa y demás, y se ha encontrado una parte en físico y otra digitalizada. Por lo tanto, acceder a algunos registros en ocasiones es bastante complejo, a excepción del SIAG, un sistema que en este momento está bien estructurado pero que sólo abarca una parte de la información necesaria.

La caracterización de los indicadores de esta base de datos facilitará muchos procesos de la institución, a partir de un sistema de información propio que irá nutriéndose a la par que crece la Universidad, por lo tanto se trata de una inversión a largo plazo.

En el presente artículo se evidencia el proceso de caracterización de los indicadores y sus etapas, que permitirá el modelado de diferentes bases de datos, con las que se recopilará la información necesaria para generar los indicadores e implementarlos, de tal manera que se acoplen adecuadamente a las bases de datos de los sistemas que se utilizan en cada una de las áreas de la institución. Las bases de datos diseñadas, además de acoplarse a las existentes, deben asegurar la integridad de los datos, lo que incluye claves principales, claves externas, restricciones, etc. Estas

características y restricciones de integridad facilitan la aplicación de reglas de negocio establecidas en la información de las tablas, lo que garantiza la precisión y fiabilidad de los datos.

II. MATERIALES Y MÉTODOS

En esta sección se presenta el marco de trabajo utilizado para el diseño de un modelo de datos para la gestión de indicadores del sistema de aseguramiento de alta calidad de la Institución Universitaria Colegio Mayor del Cauca.

El proyecto se realizó basado en el marco de trabajo ágil llamado Scrum, el cual permite a los equipos autoorganizados entregar resultados de alta calidad en tiempos cortos, para facilitar el desarrollo de las actividades y tareas planteadas al inicio del proyecto. Es un marco de trabajo para prácticas ágiles muy simple y escalable, que requiere un gran esfuerzo, ya que no se basa exclusivamente en el seguimiento de un plan sino en la adaptación continua a las circunstancias de la evolución del proyecto [1]. Su objetivo es que los miembros del equipo trabajen juntos y de forma eficiente para obtener productos complejos.

Scrum mantiene firmemente valores que ayudan a mejorar el rendimiento del trabajo en equipo, a saber: *Foco*: Enfocarse en una sola cosa a la vez, permitiendo que el trabajo en equipo ofrezca los mejores resultados posibles. *Coraje*: Sentirse apoyado por el equipo de trabajo y tener los recursos a su disposición, que da como resultado el coraje para enfrentar grandes desafíos. *Apertura*: El trabajo que se realiza en equipo permite dar a conocer las mejoras o problemas que se puedan presentar, de manera que todos manifiesten sus preocupaciones y puedan ser tenidas en cuenta. *Compromiso*: Siempre habrá el compromiso de mejorar y llegar al éxito del proyecto. *Respeto*: Es parte fundamental al trabajar en equipo, pues el compartir éxitos y fracasos demuestra en realidad el compañerismo y el respeto mutuo [2].

El ciclo de vida de SCRUM es incremental iterativo y se caracteriza por ser muy adaptable. Estas iteraciones o *sprints* son periodos que, según los casos, pueden tener duraciones desde una semana hasta no más de dos meses, en los cuales se deben realizar pequeñas entregas operativas del producto. En cada *sprint* se diseña, codifica y testea el producto con el fin de asegurar el correcto funcionamiento y permitir, además, que el cliente pueda sugerir mejoras o cambios con los cuales se tendrá mayor satisfacción con la aplicación final [3].

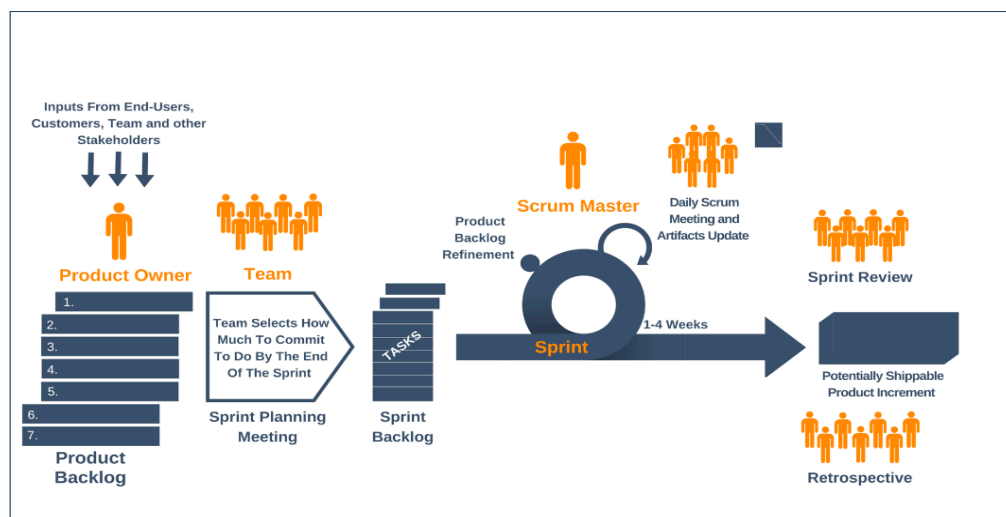


Fig. 1. Proceso de Scrum

Entre los roles que se manejan en Scrum no existe uno en específico para un gerente de proyecto, esta responsabilidad se encuentra dividida en los tres roles que conforman el equipo de Scrum [1]:

Dueño del producto (Product Owner): Aquella persona que tiene conocimiento acerca del negocio del cliente, siendo responsable del éxito del producto, pues representa a todos los interesados en el producto final. Además, es quien se responsabiliza de gestionar las expectativas de los *stakeholders* o interesados.

Scrum Master: Es el responsable del correcto funcionamiento y de garantizar el correcto entendimiento y uso de Scrum dentro de la organización, es quien ayuda al equipo a alcanzar el máximo nivel de productividad.

Equipo de desarrollo (Scrum Team): Es un equipo preparado para la construcción del producto en cuestión, con las habilidades necesarias para crear un producto de calidad.

Scrum posee una mínima cantidad de herramientas o artefactos para poder mantener organizado y llevar a cabo un proyecto. Estos artefactos son de gran apoyo debido a que ayudan a planificar y revisar cada uno de los *sprints* con el fin de aportar o mejorar el desarrollo del producto. [4] Cada uno de ellos está organizado así:

Pila del producto (Product Backlog): Es un listado que contiene las características requeridas por el cliente del producto que se va a construir y priorizadas según el orden que él determina. La pila del producto se encontrará a la vista de todos los involucrados en el proyecto. Esta es responsabilidad únicamente del Dueño del producto (*Product Owner*), quien la creará y modificará si es necesario, aunque puede recibir sugerencias o recomendaciones del Equipo [4].

Pila del sprint (Sprint Backlog): Es la recopilación sintética de ítems de la Pila del producto que se realizan al comienzo del *sprint*, negociados entre el Dueño del producto y el Equipo de desarrollo. Esta recopilación o lista se descompone con el fin de asignar una tarea de tamaño adecuado a cada uno de los integrantes del Equipo de desarrollo, de manera que se logre determinar el avance a diario.

Tablero de tareas (Scrum Taskboard): Con la lista de tareas de la Pila del *sprint* ya dividida, se procede a crear un Tablero de tareas. Un *Scrum Taskboard* se divide principalmente en tres columnas: Por hacer, En ejecución y Terminado, dentro de las cuales se colocarán las tareas respectivamente según su estado. Cada una de ellas debe visualizar una descripción de la tarea y el nombre de la persona encargada. Por otra parte, una práctica que se hace es etiquetar las tareas de manera que se logre diferenciar, por ejemplo, por prioridad, por complejidad, etc. [4].

III. RESULTADOS

En esta sección se presentan los *sprints* resultado del proceso de diseño, de igual manera se incluyen algunos artefactos como la pila de producto, la división de los *sprints* y los modelos de datos resultados en cada *sprint*. Un *sprint* es considerado una iteración de tiempo (*time-box*), aproximadamente de un mes, durante el cual se realizan cortos avances de un producto terminado, utilizable y potencialmente entregable.

A. Sprint 0

Dentro del *sprint 0* se define la visión del proyecto, los participantes, el equipo de desarrollo, el objetivo del proyecto, el objetivo del producto y el levantamiento de requisitos.

Inicialmente se establece como objetivo del proyecto caracterizar los indicadores necesarios para el aseguramiento de la calidad de la Institución Universitaria Colegio Mayor del Cauca, a través de la gestión de un sistema de información vinculado con la matriz de correlación del trabajo de grado: “Manual de Estructuración del Sistema de Aseguramiento Interno de la Calidad de la Institución Universitaria Colegio Mayor del Cauca”, y como objetivo del producto centralizar la forma de obtener información relacionada con los indicadores solicitados para el Aseguramiento Interno de la Calidad de la Institución Universitaria Colegio Mayor del Cauca. Luego del levantamiento de requerimientos se establece la pila de producto o *backlog*, que consiste en una lista ordenada según prioridad de las actividades y tareas definidas para el proyecto y que están pendientes de realizar. En la figura 2 se muestra este artefacto.

A partir de esta especificación de requerimientos se planean los *sprints* que se van a desarrollar. En la tabla 1 se presenta la planeación y distribución de los *sprints* para el desarrollo del proyecto.

ID	Actividad	Descripción	Tareas
A1	Reconocer e interpretar las características e indicadores.	Interpretación y reconocimiento de las características e indicadores propuestos desde la actividad académica adelantada para realizar el manual existente.	T1: Examinar las características e indicadores del manual. T2: Interpretar las características e indicadores consolidados en el manual. T3: Revisión de dudas de las características e indicadores del manual.
A2	Establecer categorías de agrupación para los indicadores.	Establecer categorías de agrupación para los indicadores, en función de los actores de las diferentes áreas que desempeñan roles de investigación, bienestar o impacto social.	T4: Identificar los actores que hacen parte de los procesos institucionales. T5: Establecer las categorías de agrupación para los indicadores.
A3	Obtener el total de los indicadores y su respectiva información.	Obtener el universo de los indicadores fundamentales para realizar el análisis y seguimiento de los procesos vinculados a autoevaluación y calidad, además la información que cada uno de estos necesita.	T6: Estructurar el archivo de consolidación de los indicadores. T7: Investigar y analizar los indicadores requeridos de alta calidad. T8: Consolidar los indicadores finales. T9: Revisión de los indicadores finales con vicerrectoría.
A4	Realizar un diagnóstico de las bases de datos existentes.	Efectuar un diagnóstico de las bases de datos existentes de los sistemas de información académicos actuales.	T10: Revisión de los indicadores con jefe de desarrollo de la institución. T11: Obtener la información existente de las bases de datos de los sistemas de la institución. T12: Identificar la información faltante en las bases de datos.
A5	Integrar y relacionar los datos.	Integrar y relacionar la información de los indicadores consolidados con respecto a la información de los sistemas de información legados de la institución.	T13: Consolidar la información existente y la información faltante de las bases de datos.
A6	Diseñar la base de datos	Diseñar una base de datos relacional para gestionar de manera eficiente los indicadores consolidados.	T14: Normalizar las nuevas tablas de acuerdo con las bases de datos existentes. T15: Identificar y establecer relaciones de las tablas de la base de datos consolidada. T16: Diseño final de la base de datos relacional para los indicadores.
A7	Diagramar la base de datos.	Modelar y realizar el diagrama de la base de datos que permita gestionar los indicadores.	T17: Seleccionar e instalar la herramienta para realizar el diagrama de la base de datos. T18: Modelar y realizar el diagrama de la base de datos.
A8	Implementar la base de datos	Realizar la implementación del modelo de la base de datos garantizando la articulación con el SIAG.	T19: Crear las nuevas tablas y campos en las bases de datos existentes. T20: Crear las relaciones entre las tablas de la base de datos.

Fig. 2. Pila de producto

Tabla 1. Planificación de *sprints*

Sprint	Actividad	Est.
1	Reconocer e interpretar las características e indicadores	6 días
	T1: Examinar las características e indicadores del manual	2 días
	T2: Interpretar las características e indicadores consolidados en el manual.	3 días

Sprint	Actividad	Est.
1	T3: Revisión de dudas de las características e indicadores del manual	1 día
	Establecer categorías de agrupación para los indicadores	9 días
	T4: Identificar los actores que hacen parte de los procesos institucionales	1 día
2	T5: Establecer las categorías de agrupación para los indicadores	8 días
	Obtener el total de los indicadores	33 días
	T6: Estructurar el archivo de consolidación de los indicadores	1 día
	T7: Investigar y analizar los indicadores requeridos de alta calidad, extrayendo la información necesaria para cada uno de ellos.	6 días
3	T8: Consolidar los indicadores finales	25 días
	T9: Revisión de los indicadores finales con vicerrectoría	1 día
	Realizar un diagnóstico de las bases de datos existentes	7 días
	T10: Revisión de los indicadores con el jefe de desarrollo de la institución	1 día
4	T11: Obtener la información existente de las bases de datos de los sistemas de la institución	4 días
	T12: Identificar la información faltante en las bases de datos	2 días
	Integrar y relacionar los datos	3 días
	T13: Consolidar la información existente y la información faltante de las bases de datos	3 días
5	Diseñar la base de datos	5 días
	T14: Normalizar las nuevas tablas de acuerdo con las bases de datos existentes	1 día
	T15: Identificar y establecer relaciones de las tablas de la base de datos consolidada	2 días
	T16: Diseño final de la base de datos relacional para los indicadores	2 días
	Diagramar la base de datos.	5 días
	T17: Seleccionar e instalar la herramienta para realizar el diagrama de la base de datos.	1 día
6	T18: Modelar y realizar el diagrama de la base de datos	4 días

Sprint	Actividad	Est.
5	Implementar la base de datos	3 días
	T19: Crear las nuevas tablas y campos en las bases de datos existentes	2 días
	T20: Crear las relaciones entre las tablas de la base de datos	1 día

Fuente: Elaboración propia.

Para el seguimiento de las tareas se implementó la metodología “Kanban” y la herramienta “Trello”, en la cual se pueden registrar y asignar las tareas definidas. El tablero se divide en tres fases principales: “Por hacer”, “En ejecución” y “Terminado”. La herramienta permite hacer seguimiento de tiempo y control de cambios, lo cual fue significativo en el desarrollo e implementación del proyecto.

B. Sprint 1

En el transcurso del primer *sprint* se realizaron las tareas de acuerdo con los tiempos estimados inicialmente; la tarea 1 se cumplió en un menor tiempo del planeado. Las categorías de agrupación se establecieron de la siguiente manera, teniendo en cuenta los actores de las diferentes áreas que desenvuelven roles en la gestión de la alta calidad de la institución [5] [6]. En la figura 3 se presenta el tablero Kanban al final del Sprint 1. Las tareas ejecutadas en este *sprint* están etiquetadas en el tablero con color rosa (ver figura 3).

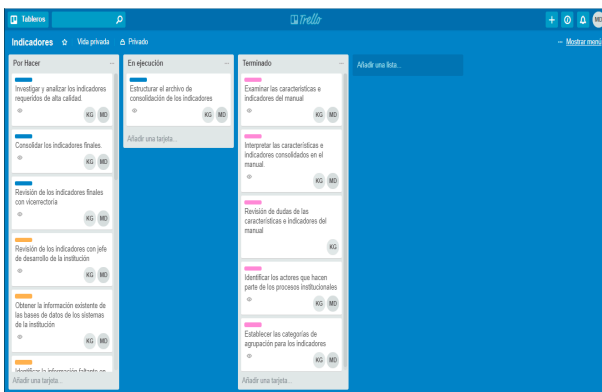


Fig. 3. Tablero Kanban correspondiente al Sprint 1

C. Sprint 2

A partir de las categorías de agrupación definidas en el *sprint* anterior y de la matriz de correlación del manual de estructuración de aseguramiento de la calidad [7] de la institución, se identificaron los indicadores finales vinculados a los procesos de alta calidad, se realiza un bosquejo de la presentación de la información identificada para cada uno de los indicadores, y el modelo de la interfaz gráfica, según las categorías de agrupación y los indicadores identificados. Durante este proceso de obtención del universo de indicadores se emplearon tres días más de lo planeado, dado que la cantidad de indicadores es considerable, y además se presentaron pequeños inconvenientes de tiempo en la investigación y el análisis, pero se lograron solucionar rápidamente, lo que permitió avanzar de manera favorable con el siguiente *sprint*. Las actividades correspondientes a esta iteración están etiquetadas en la tabla con color azul (ver figura 4).

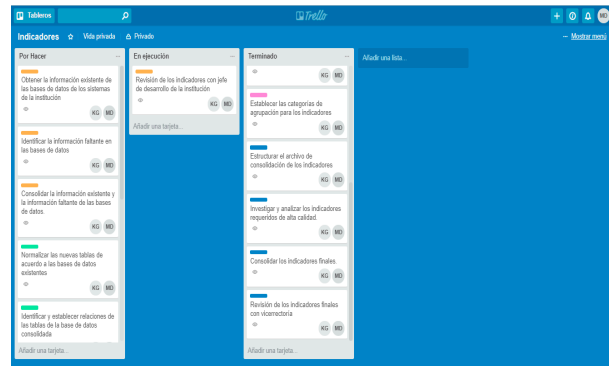


Fig. 4. Tablero Kanban correspondiente al Sprint 2

D. Sprint 3

Una vez que se obtuvo el total de los indicadores, se realizó un diagnóstico a las bases de datos de los sistemas de información existentes en la institución, para integrarlo posteriormente con la información identificada. A partir de este diagnóstico se incluyeron en las tablas algunos atributos o campos que se hicieron necesarios para cumplir con el objetivo del producto y los requerimientos [8].

Las tareas del tercer *sprint* se ejecutaron sin ningún inconveniente, durante los tiempos establecidos. La

segunda tarea se cumplió en menor tiempo del que se estimó al inicio, de esta manera el *sprint* culmina de forma exitosa. Las actividades correspondientes a esta iteración están etiquetadas en la tabla con color naranja (ver figura 5).

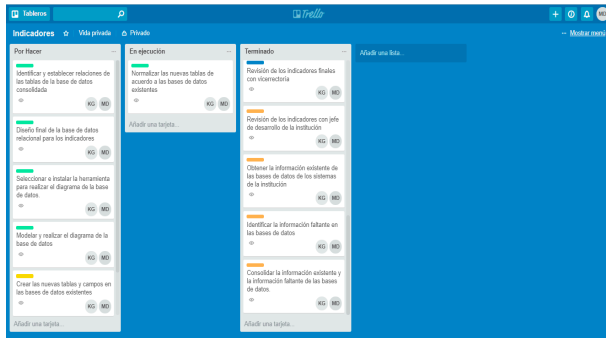


Fig. 5. Tablero Kanban correspondiente al Sprint 3

E. Sprint 4

Se realizaron los modelos de datos [9] correspondientes (SIAG, ADMISIONES, BIENESTAR EX-TENSIÓN), con la información integrada, tanto del sistema de información ligado con la obtenida de los indicadores; como resultado se obtuvieron los modelos en formato PNG (imagen) y formato xml (DB-Designer). Las actividades correspondientes a esta iteración están etiquetadas en la tabla con color naranja (ver figura 6).

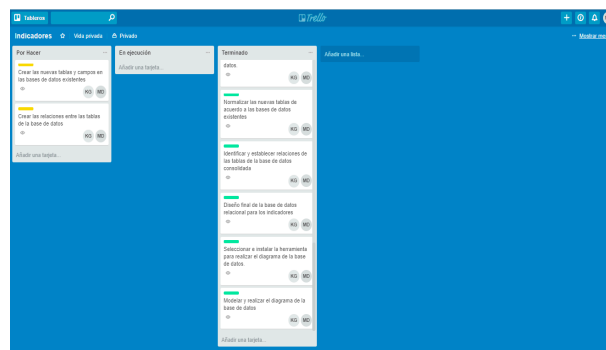


Fig. 6. Tablero Kanban correspondiente al Sprint 4

F. Sprint 5

A momento de realizar la implementación de las nuevas tablas y campos, se hizo con respecto a las bases de datos del 04 de Diciembre de 2017, se en-

tregan los *scripts* correspondientes para las bases de datos SIAG, ADMISIONES, BIENESTAR y EX-TENSIÓN. El quinto *sprint* termina de manera exitosa, se cumplieron satisfactoriamente las tareas que se programaron y la creación de las tablas y campos en el motor de base de datos se realizó en un día menos de lo estimado, de esta manera termina el ultimo *sprint* del proyecto. Para esta iteración las actividades desarrolladas se etiquetan con color verde (ver figura 7).

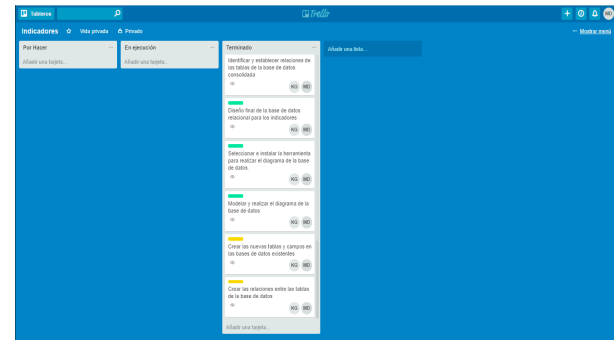


Fig. 7. Tablero Kanban correspondiente al Sprint 5

IV. DISCUSIÓN

Inicialmente se obtuvo una agrupación para los indicadores estableciendo las categorías a partir de la información referente a los actores de las diferentes áreas que desenvuelven roles en la gestión de la alta calidad de la institución [10]. La agrupación establecida se muestra en la tabla 2

Tabla 2. Categorías definidas para los indicadores

Id	Categoría
1	Docentes
2	Investigación
3	Recursos Físicos
4	Recursos Informáticos
5	Programas Académicos
6	Matrículas
7	Estímulos Económicos
8	Inscritos
9	Egresados
10	Bienestar Universitario
11	Financiación
12	Plan De Estudios

Id	Categoría
13	Movilización Y Comisión De Estudios
14	Convenios
15	Representantes
16	Proyección Social
17	Icfes
18	Recursos Bibliográficos

Fuente: Elaboración propia.

A partir de las categorías de agrupación definidas y la matriz de correlación del manual de estructuración de aseguramiento de la calidad de la institución, se identificaron los indicadores finales vinculados a los procesos de alta calidad [7]. En la tabla 3 vemos el número de indicadores identificados por categoría.

Tabla 3. Número de indicadores identificados por categoría

Id	Categoría	N. indicadores identificados
1	Indicadores Docentes	22
2	Indicadores Investigación	24
3	Indicadores Recursos Físicos	7
4	Indicadores Recursos Informáticos	6
5	Indicadores Programas Académicos	10
6	Indicadores Matrículas	33
7	Indicadores Estímulos Económicos	34
8	Indicadores Inscritos	23
9	Indicadores Egresados	29
10	Indicadores Bienestar Universitario	19
11	Indicadores De Financiación	3
12	Indicadores Plan De Estudios	2
13	Indicadores Movilización Y Comisión De Estudios	46
14	Indicadores Convenios	6
15	Indicadores Representantes	10
16	Indicadores Proyección Social	10
17	Indicadores Icfes	4
18	Indicadores Recursos Bibliográficos	6

Fuente: Elaboración propia.

Finalmente, luego del desarrollo del proyecto, se obtuvieron los modelos de bases de datos integrados. En la figura 7 se presenta el modelo integrado de admisiones, modificado y adaptado para cumplir con el objetivo del producto.

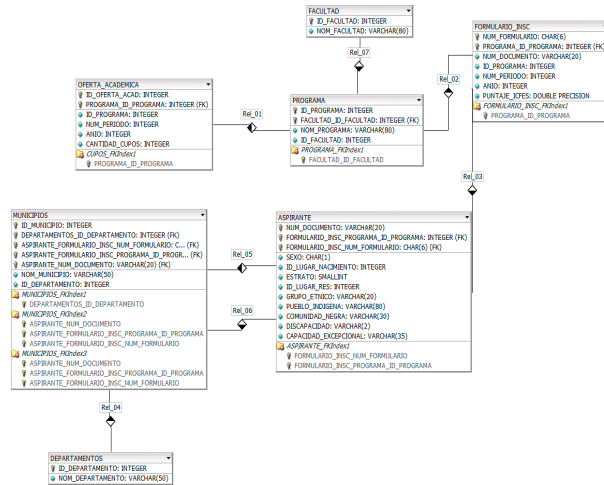


Fig. 6. Modelo de base de datos adaptado para el proceso de admisiones

V. CONCLUSIONES

Los indicadores de gestión de alta calidad son elementos que reflejan los resultados del trabajo de cada una de las áreas de la institución y el cumplimiento de metas y objetivos, los mismos que permiten alcanzar la acreditación de alta calidad para la institución y sus programas.

Con el fin de realizar un seguimiento al proceso de autoevaluación con fines de acreditación de alta calidad, los indicadores deben ser clasificados en diferentes áreas estratégicas, pero con información clara y fácil de recolectar.

Es necesario centralizar la información bajo un modelo de datos que permita facilitar los procesos de evaluación y estructuración de informes, con el fin de evitar información dispersa y no uniforme, resultados diferentes y saturación de las unidades productoras.

REFERENCIAS

- [1] M. Alaimo, “Introducción a la agilidad y Scrum”, 2013. [En línea]. Disponible en: <http://www.martinalaimo.com/es/blog/enloqueando-2015-09-mas-alla-del-producto-funcionando>.
- [2] L. De Seta, “Todo Scrum en 1 sola página”, 2013. [En línea]. Disponible en: <https://dosideas.com/noticias/metodologias/981-scrum-en-1-sola-pagina>.
- [3] J. Palacio, “Flexibilidad con Scrum: Principios de diseño e implantación de Campos de Scrum”, 2007. [En línea]. Disponible en: <https://www.safe-creative.org/work/0710210187520>.
- [4] K. Schwaber y J. Sutherland, “Guía de SCRUM”, 2009. [En línea]. Disponible en: <http://www.scrum.com>: <http://www.scrum.org>.
- [5] J. Martínez, Manual de estructuración del sistema de aseguramiento interno de la calidad de la Institución Universitaria Colegio Mayor del Cauca, Popayán: Institución Universitaria Colegio Mayor del Cauca (Trabajo del grado), 2017.
- [6] Ministerio de Educación Nacional - Consejo Nacional de Acreditación, “Acreditación de Programas de Pregrado”, 2014. [En línea]. Disponible en: <https://www.cna.gov.co/1741/article-186377.html>.
- [7] Ministerio de Educación Nacional (MEN) - Sistema de Aseguramiento de la Calidad (SACES), “¿Qué es aseguramiento de la calidad?”, s. f. [En línea]. Disponible en: <http://www.mineducacion.gov.co/sistemasdeinformacion/1735/article-227110.html>.
- [8] C. Quintero, “Modelos de Datos”, 2013. [En línea]. Disponible en: <http://carmenq27.blogspot.com.co/2013/09/modelos-de-datos-introduccion-modelo-y.html>.
- [9] Universidad San Tomás, “Lenguaje de Modelado”, s. f. [En línea]. Disponible en: http://soda.ustadistancia.edu.co/enlinea/mariocontreras_lenguajedemodelaje/.
- [10] Universidad Central, “¿Qué es la acreditación?”, s. f. [En línea]. Disponible en: <http://www.uccentral.edu.co/la-universidad/vicerrectoria-academica/acreditacion/que-es-la-acreditacion>.

Capítulo VI

Agrupamiento Jerárquico sobre Uso de las TIC en Colombia: Un caso de Aplicación del *Machine Learning*

Víctor Daniel Gil Vera - victor.gilve@amigo.edu.co

Catalina Quintero López - catalina.quintero@amigo.edu.co

Vicerrectoría de Investigaciones, Universidad Católica Luis Amigó, Medellín, Colombia

I. INTRODUCCIÓN

Tanto para los países desarrollados como para los países en desarrollo, las TIC son un factor clave para mejorar la eficiencia productiva, el crecimiento económico y el desarrollo social [1]. Debido al cambio de paradigma tecno-productivo y al crecimiento de las TIC en el mundo en las dos últimas décadas del siglo XXI, vale la pena hacer un análisis sobre el uso que hacen de las TIC las empresas pertenecientes a los tres principales sectores de la economía colombiana: manufactura, comercio y servicios.

En Colombia, a pesar de que la mayoría de empresas manufactureras, comerciales y de servicios están presentes en la web, los niveles de inversión en TIC siguen siendo bajos [2]. El factor más importante que limita la digitalización de los empleados en Colombia es la falta de educación [2]. Es importante conocer los impulsores del uso de las TIC, ya que los encargados de formular políticas en el país han emitido planes para cerrar la brecha digital que ha surgido con el advenimiento de estas tecnologías [3]. La educación y los ingresos son dos de los factores más importantes para explicar la brecha en el uso de las cuatro TIC principales en el país: telefonía fija y móvil, Internet y SMS. Colombia es uno de los pocos países de América Latina donde también existe una brecha digital por razones de género [2].

Los dirigentes de las empresas en Colombia deben implementar cursos de capacitación que se ajusten a los niveles de formación de los emplea-

dos. También se les debe dar mayor importancia y apoyo a los programas gubernamentales y a las organizaciones no gubernamentales (ONG) que educan y capacitan a las personas de bajos ingresos en herramientas de TIC. Características como las condiciones geográficas, el alto costo de la infraestructura y los bajos niveles de educación requieren mayores niveles de inversión. Esto es importante, ya que internet y los servicios móviles amplían la brecha entre pobres y no pobres cuando el conocimiento básico está ausente [3]. Actualmente, muchos dispositivos de telefonía móvil ofrecen diversas funciones informáticas, lo que hace que la disponibilidad de una computadora sea una barrera menor para acceder a internet [3]. El uso masificado de internet en sectores económicos como la banca y el comercio no será posible para las poblaciones de bajos ingresos hasta que existan condiciones para su uso [3]. Es importante diseñar programas para sectores desfavorecidos de la población que deben ir más allá del subsidio de dispositivos móviles y de computadoras, como ha sido frecuente en los países latinoamericanos.

Considerando todo lo anterior, el objetivo de este trabajo fue realizar una clasificación y un análisis del uso que las empresas en Colombia hacen de las TIC empleando aprendizaje de máquinas, específicamente el agrupamiento jerárquico. La pregunta de investigación considerada en este trabajo fue:

PI. ¿Cómo se agrupan las empresas en Colombia según el uso que hacen de las TIC y qué se puede concluir de ello?

En el análisis se consideraron empresas manufactureras, comerciales y de servicios. Se empleó la base de datos de indicadores sobre la posesión y uso de las TIC en empresas colombianas [4]. Se utilizó el algoritmo de agrupación jerárquico, el cual consiste en un método de análisis que se fundamenta en la formación de grupos según las características de similitud de los datos.

Se concluye que las empresas de servicios son las que mayor uso hacen de las TIC. Por el contrario, el personal que trabaja en empresas manufactureras es el que menos uso hace de ellas, situación que puede obedecer a que la mayoría no sabe cómo aplicarlas o no emplean tecnología en el desarrollo de sus actividades [5]. Es necesario que estas empresas realicen un estudio profundo sobre sus necesidades reales para que puedan ser competentes en el mercado y no destruyan valor.

II. LAS TIC EN COLOMBIA

Las TIC son una tecnología heterogénea que permite almacenar información, procesarla y darla a conocer de diferentes maneras [6]. En las dos últimas décadas se ha producido en todos los sectores de la economía una auténtica explosión de las TIC, debido al uso masivo de ordenadores personales y a la expansión del internet, lo que ha facilitado el acceso de las personas a una creciente cantidad de información [7], [8].

En el sector servicios de Colombia el uso de las TIC tiene una gran importancia. Existe una correlación positiva y directa entre capital TIC y las aplicaciones informáticas en la productividad laboral de los servicios [9]together with non-ICT capital, training, and the proportion of foreign capital and permanent workers are the main determinants of productivity increase in the Colombian services firms. (English. Las inversiones en recursos tecnológicos de TIC, junto con el capital no TIC, la capacitación, el porcentaje de capital extranjero y de empleados permanentes son los recursos más importantes asociados positivamente con la productividad de las empresas de servicios de Colombia [9]together with non-ICT capital, training, and the proportion of foreign capital and permanent

workers are the main determinants of productivity increase in the Colombian services firms. (English.

El Gobierno colombiano en los últimos años ha enfatizado en la difusión y uso de las TIC por parte de la ciudadanía en general y de las empresas del sector productivo en particular [9]together with non-ICT capital, training, and the proportion of foreign capital and permanent workers are the main determinants of productivity increase in the Colombian services firms. (English. El Plan Nacional TIC de 2008 al igual que el Plan Vive Digital han pretendido masificar el uso de las TIC entre el empresariado colombiano, en particular las Pymes, para lo cual han propuesto algunas medidas de política [9]together with non-ICT capital, training, and the proportion of foreign capital and permanent workers are the main determinants of productivity increase in the Colombian services firms. (English. Las empresas que invierten en TIC obtienen beneficios en productividad en comparación con las que no invierten. Se hace necesario entonces que el Gobierno implemente medidas más audaces para incentivar a las Pymes a invertir en estas tecnologías [9]together with non-ICT capital, training, and the proportion of foreign capital and permanent workers are the main determinants of productivity increase in the Colombian services firms. (English.

Una fuerza laboral más capacitada coadyuva a mejoras en productividad. Medidas de política han de ponerse en marcha usando para ello al SENA, y creando incentivos (tributarios y financieros) para ayudar a las Pymes y en especial a los microempresarios a capacitar su personal de trabajo [9]together with non-ICT capital, training, and the proportion of foreign capital and permanent workers are the main determinants of productivity increase in the Colombian services firms. (English. Así, por ejemplo, las principales TIC utilizadas en los hoteles incluyen la dotación de hardware y software de gestión, tecnologías de conexión a redes, aplicaciones de mercadeo electrónico y de ventas, lo cual impacta positivamente en la expansión del mercado, la imagen y la calidad y además genera una ventaja competitiva, sin embargo, no les están sacando todo el potencial que se puede aprovechar de las TIC [10]que ha conllevado cambios en la oferta, la demanda, las

preferencias de los clientes, así como en la rapidez en el desarrollo y la propagación de las tecnologías de la información y de la comunicación (TIC).

Las TIC pueden ayudar a las empresas a mejorar su eficiencia y a ser más competitivas, no obstante, este proceso debe estar asistido por un esfuerzo de planeación, formación a los empleados en el uso y adopción de las TIC y cambio organizacional. Así, por ejemplo, una de las principales barreras para que los clientes de los hoteles utilicen las TIC es la percepción sobre la baja calidad del servicio, la falta de información, los problemas logísticos en el cumplimiento de lo contratado y la inseguridad en pagos vía electrónica [10] que ha conllevado cambios en la oferta, la demanda, las preferencias de los clientes, así como en la rapidez en el desarrollo y la propagación de las tecnologías de la información y de la comunicación (TIC). La implementación de nuevas tecnologías o TIC ha revolucionado la productividad de las grandes empresas, pero también es posible aplicarlas en las Pymes [11]. Las empresas desde muy pequeñas deben buscar estandarizar sus actividades y eliminar en la medida de lo posible los procesos manuales. El hecho de que las empresas tengan pocos empleados no significa que deban utilizar su tiempo en tareas administrativas de manera indiscriminada, sin antes pensar qué se podría hacer mejor y más fácil, desde escribir un proceso cuidadosamente en un manual hasta desarrollar bases de datos que generen facturas automáticas [11]. La Tabla 1 presenta algunas de las aplicaciones más utilizadas en la actualidad en las empresas colombianas:

Tabla 1. Aplicaciones y usos

Aplicación	Función
Mercado pago/ Paypal	Realizar transacciones con clientes de todo el mundo
Skype / Hangouts	Realizar videoconferencias
WhatsApp	Facilitar el contacto con los clientes en vez de utilizar el correo o el teléfono
Youtube /Vine / Wideo	Hacer y promocionar videos corporativos

Aplicación	Función
Mercado Libre / e-Bay/Amazon y Alibaba	Comprar y vender
Linkedin/ Face- book/Google Ads /blogs	Informar a los clientes, mantener una comunicación periódica y promocionar sus productos y servicios
Mailchimp	Enviar newsletters
Trello / Google Calendar	Organizar eventos y tareas

Fuente: Elaboración propia.

En la actualidad, existen muchos desafíos para las empresas manufactureras, comerciales y de servicios en el país. La Tabla 2 presenta la descripción de algunos de ellos en cada uno de los tres sectores de la economía.

Tabla 2. Desafíos

Sector	Desafíos
Manufactura	Automatización de la producción. Para ello se requiere establecer iniciativas gubernamentales, capacitar el recurso humano con las competencias que estas nuevas tecnologías requieren, formar consultores especializados y mejorar la ciberseguridad.
Comercio	<ul style="list-style-type: none"> - Diferentes y mejores formas de organización y coordinación. - Costes menores de procesamiento de información. - Posibilidad de ofrecer servicios nuevos. - Redefinición de los segmentos de mercado. - Reformulación de las actividades tradicionales. - Mayor grado de cooperación. - Utilización de nuevas herramientas competitivas. - Modificación de las pautas tradicionales de rivalidad competitiva. - Nuevos hábitos y comportamientos de compra. - Demanda de nuevos servicios. - Aparición de nuevos formatos comerciales.

Sector	Desafíos
Servicios	<ul style="list-style-type: none"> - Creación de herramientas digitales que puedan disminuir el tiempo de los procesos para la generación de trámites de registros, teniendo en cuenta que la información debe estar almacenada en un mismo lugar y a partir de datos existentes. - Creación de herramientas tecnológicas que incidan en la disminución del índice de accidentalidad de operarios y puedan ser utilizadas en los procesos de inducción y entrenamiento. - Creación de herramientas tecnológicas que gestionen la movilidad de pasajeros que se movilizan en medios no masivos de transporte. - Creación de herramientas tecnológicas que permitan la validación y formación de precios de los productos con los que trabajan los cultivadores y compradores, compartiendo los datos de las transacciones que realizan entre ellos.

Fuente: Elaboración propia.

El crecimiento de las economías está influenciado por la inversión en TIC, en innovación y en educación [12]. Tanto las economías de países desarrollados como las de países emergentes, tienen en estos tres elementos sus principales factores de crecimiento, y los países que se destacan son aquellos que han mantenido una estrategia coherente y de largo plazo. Obviamente son las empresas las que, en última instancia, mueven las economías y es en ellas donde se soporta el resultado [12].

Las empresas que no usan las TIC para atender las exigencias del mercado moderno ceden terreno. Muchas de ellas pierden oportunidades de negocios por no contar con software adecuado. Las empresas que no utilizan TIC como herramientas competitivas en sus negocios, limitan su crecimiento empresarial y pueden perder ingresos [12].

En Colombia, a pesar de los importantes progresos en penetración de internet, si se comparan

las inversiones en TIC entre las empresas grandes frente a las Pymes, la brecha es de proporciones gigantescas [12]. Colombia es un país donde el 98 % de las empresas son Pymes y de este porcentaje solo 2 %, que corresponde a las empresas medianas, están preparadas para competir. Y aunque más de 80 % de las Pymes invierten en computadores, sólo 55 % tienen página web y apenas un 33 % invierten en sistemas de información de gestión [12].

De lo anterior, se deduce que muy pocas de ellas tienen dentro de su estrategia la adopción de TIC como un factor de competitividad y crecimiento y cada día que pase sin que haya conciencia en los empresarios sobre esta situación, significa ampliar la brecha [12]. Si una empresa que adopta TIC crece 10 % sus ventas y una que no, pierde hasta el 30 % de sus ingresos, estamos frente a un efecto devastador para aquellos que no lo asuman con la seriedad que esto implica [12].

Los beneficios de adquirir TIC se dan cuando se obtiene conciencia de que las ventajas competitivas y la sostenibilidad de los negocios se aseguran con el adecuado uso de la tecnología. Invertir en TIC es un proceso que, una vez iniciado, no termina, es complejo e implica determinación y convicción por parte de los accionistas y directivos de las empresas [12]. Debe darse paso a paso, iniciando con la apropiación de infraestructura y software genérico, hasta llegar a soluciones de software especializado. La idea fundamental con la adquisición de estas infraestructuras es que las empresas sean competitivas y se preparen para asegurar un futuro.

El uso de las TIC sumado a factores financieros varía en importancia entre las diferentes industrias con diferentes niveles de desarrollo tecnológico [13]. Las industrias en desarrollo que cuentan con una infraestructura tecnológica pobre dependen de los préstamos y las inversiones de entidades financieras [13]. Las industrias desarrolladas, por el contrario, cuentan con recursos para innovar y aplicar tecnología de manera permanente, lo que las hace más competitivas. La Figura 1 presenta el uso global de las TIC en el período 2001-2018 por cada 100 habitantes.

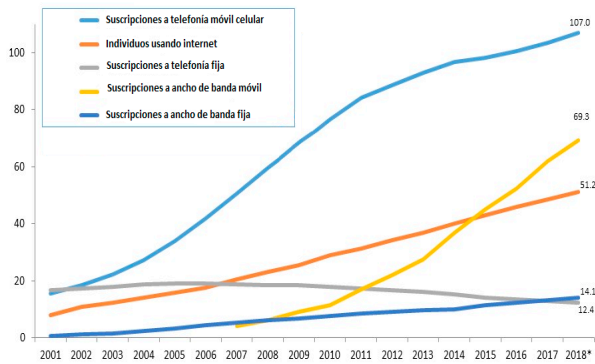


Fig. 1. Uso de las TIC en el mundo.

Fuente: Elaboración de los autores con base en [14]

Los principales sectores que conforman la economía colombiana son:

Manufacturero: Este sector representa una piedra angular de muchas economías nacionales, y constituye un sector crucial para la generación de empleos productivos, el cambio estructural y el crecimiento económico sostenible [15]. Este sector es el cuarto más representativo de la actividad productiva en la economía colombiana [16]. Las principales ventajas de este sector son: las preferencias arancelarias sobre varios productos en diversos mercados a través de los acuerdos comerciales firmados por Colombia, la aplicación de Buenas Prácticas de Fabricación (BPF) en las etapas del proceso de producción, el alto nivel de calidad e innovación en productos, con certificaciones reconocidas a nivel internacional y la flexibilidad en la producción para adaptarse a las demandas del mercado internacional [17]. La Tabla 1 presenta los principales países que importan productos manufactureros de Colombia.

Tabla 3. Destinos principales de exportaciones manufactureras

País	USD millones	Porcentaje
Ecuador	1.060.000	15,2 %
EE. UU.	829,7	11,8 %
Perú	753	10,7 %
Venezuela	719,2	10,2 %
Brasil	660,3	9,4 %

Fuente: Elaboración propia.

Comercio: En Colombia, el sector terciario representa aproximadamente el 50 % del producto, y específicamente el comercio es una de las actividades que más contribuyen a la actividad económica nacional. En términos de empleo, este sector representó aproximadamente el 11 % en el año 1970 y un 25,4 % en el año 1995 [18]. El crecimiento de la economía colombiana en 2018 fue de 2,7 %, jalado por la actividad pública y el comercio [19].

Servicios: Este sector representa el 57,5 % del PIB nacional, ofrece un gran potencial de crecimiento y se convierte en una de las principales garantías comerciales para el país. Actualmente, se están llevando a cabo iniciativas para explorar nuevas oportunidades y mercados, mejorar la política del sector y ampliar la oferta nacional. La base de datos de “Indicadores básicos de posesión y uso de las tecnologías de la información y la comunicación TIC”, desarrollada por la Dirección Administrativa Nacional de Estadística (DANE), presenta un resumen de los principales usos de las TIC en empresas colombianas (ver Tabla 2).

Tabla 4. Uso de las TIC en empresas colombianas

	Empresas con computador (%)	Empresas que usan internet (%)	Empresas con página web (%)	Personal que usa computador (%)	Personal que usa internet (%)
Manufactura	99,3	99,3	71,8	51,6	50,3
Comercio	99,4	99,4	60,1	67,2	64,4
Servicios					
<i>E-mail y almacenamiento</i>	98,8	99,5	80,8	68,2	68,1

	Empresas con computador (%)	Empresas que usan internet (%)	Empresas con página web (%)	Personal que usa computador (%)	Personal que usa internet (%)
<i>Alojamiento/vivienda</i>	100,0	100,0	89,5	62,3	62,1
<i>Restaurantes y bares</i>	99,5	99,5	67,0	47,7	46,7
<i>Actividades de publicación</i>	100,0	100,0	92,1	88,0	86,5
<i>Producción de películas</i>	97,3	100,0	89,2	97,7	85,0
<i>Programación y transmisión de TV</i>	100,0	100,0	88,1	90,9	94,6
<i>Telecomunicaciones</i>	100,0	100,0	78,1	88,5	89,9
<i>Desarrollo y sistemas de información</i>	100,0	100,0	96,2	98,1	95,4
<i>Actividades inmobiliarias</i>	99,5	100,0	74,5	71,9	70,8
<i>Profesionales, científicos y técnicos</i>	99,9	99,7	84,9	84,2	84,5
<i>Agencias de viaje</i>	100,0	98,7	92,3	97,2	97,5
<i>Investigadores privados y seguridad</i>	98,9	99,3	63,7	33,0	32,9
<i>Oficinas administrativas y de soporte</i>	98,8	100,0	80,2	88,9	79,5
<i>Educación superior privada</i>	100,0	100,0	99,4	97,5	97,8
<i>Salud humana privada</i>	100,0	100,0	78,5	88,4	82,3
<i>Juegos de azar, actividades deportivas y recreativas</i>	100,0	100,0	84,0	75,5	64,6
<i>Otros servicios y actividades</i>	100,0	100,0	81,4	66,2	69,1

Fuente: Tomado de [4].

III. APRENDIZAJE DE MÁQUINAS

El aprendizaje de máquinas (*Machine Learning*) es una de las tecnologías emergentes que ha captado la atención de académicos y empresarios, y se espera que evolucione en un futuro cercano [20]. Es también considerado como una forma de inteligencia artificial [21], en la que los algoritmos realizan predicciones para interpretar datos y “aprender”, sin instrucciones estáticas [22].

En la actualidad existen dos formas principales de aprendizaje de máquinas: supervisado y no supervisado. En el aprendizaje supervisado los algoritmos reciben datos de entrenamiento, que se analizan en busca de características importantes para la clasificación y el etiquetado [23]. El modelo se “entrena” con estos datos antes de ser probado con datos sin etiquetar. El no supervisado se utiliza para identificar patrones sin entrenamiento. Las formas

comunes son análisis de conglomerados (los datos se agrupan por patrones de características) o asociación (las reglas se descubren mediante el tratamiento de datos) [23].

Tabla 5. Clasificación de las técnicas de *Machine Learning*

<i>Aprendizaje de máquinas</i>	<i>Supervisado:</i> los datos están pre-categorizados o son numéricos	<i>Clasificación:</i> predice una categoría
		<i>Regresión:</i> predice un número
	<i>No supervisado:</i> los datos no están etiquetados en ninguna forma	<i>Agrupación:</i> divide por patrones similares
		<i>Asociación:</i> identifica secuencias
		<i>Reducción de dimensiones:</i> encuentra dependencias ocultas

Fuente: Elaboración propia.

A. Agrupamiento jerárquico

La agrupación jerárquica es muy popular en dominios como las ciencias sociales y la biología. El motivo de su popularidad se puede atribuir al hecho de que, a diferencia de *k-means* u otras técnicas de agrupamiento, permite comprender los datos en diferentes niveles de granularidad [24]. La agrupación jerárquica permite particionar recursivamente n puntos de datos en 2, 3, 4,... o n grupos. Un agrupamiento jerárquico puede representarse como un árbol arraigado (dendrograma) donde todas las hojas corresponden a los puntos de datos dados. Cada nodo interno representa un grupo que consta de los puntos de datos correspondientes a sus hojas descendientes (ver Figura 2).

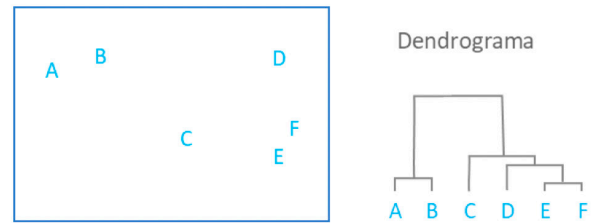


Fig. 2. Dendrograma

Fuente: Adaptado de [25].

En la revisión del estado del arte, diferentes investigadores han empleado esta técnica de agrupación. Así, por ejemplo, en [26] la emplearon para agrupar las diferentes causas de muerte en Colombia. En [27] la emplearon en la identificación óptima de módulos en sistemas mecánicos, eléctricos y de plomería (MEP). En [28] proponen un método para encontrar conjuntos de genes coexpresados, basados en índices de validación de conglomerados como una medida de similitud para grupos de genes individuales, y una combinación de variantes de agrupamiento jerárquico para generar grupos candidatos.

IV. METODOLOGÍA

En el análisis de datos se empleó la información presentada en la Tabla 2 (Indicadores básicos de posesión y uso de las TIC en empresas colombianas) y el software estadístico IBM SPSS 25, el enlace entre grupos y la función de disimilitud “Distancia Euclídea al Cuadrado”. El cálculo de la distancia euclídea al cuadrado es un factor importante en la mayoría de métodos de aprendizaje de máquinas (ver ecuación 1).

$$D_{ij} = \sqrt{\sum_{k=1}^n (x_{ki} - x_{kj})^2} \quad (1)$$

Donde:

D_{ij} es la distancia entre los casos i y j
 x_{kj} valor de la variable para el caso j

IBM SPSS 25 permite realizar agrupaciones jerárquicas desde la barra de menú: Analizar - Clasificar - Clúster jerárquico (ver Figura 3).

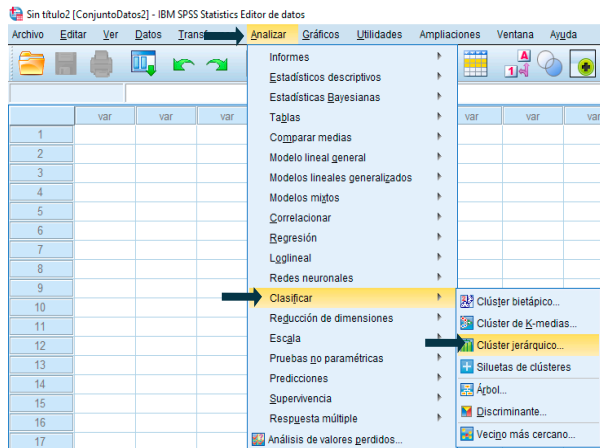


Fig. 3. Agrupación jerárquica SPSS 25

Fuente: Elaboración propia.

V. RESULTADOS

La Tabla 6 y la Figura 4 presentan los resultados de la conglomeración de los datos. Como se mencionó anteriormente, se empleó la Distancia euclídea al cuadrado y el enlace promedio entre grupos. En la Figura 3, se observa que el primer nivel está conformado en su mayoría por empresas de servicios (restaurantes y bares, producción de películas, producción y transmisión de TV, agencias de viaje, educación superior privada, empresas que realizan actividades de publicación, desarrollo y sistemas de información, telecomunicaciones, profesionales científicos y técnicos, y empresas de salud privadas) y por empresas manufactureras.

La mayoría de estas empresas usan computador e internet y cuentan con una página web; igualmente sus empleados hacen uso de estas herramientas.

El segundo nivel está conformado por empresas de servicios (desarrollo y sistemas de información, telecomunicaciones, profesionales científicos y técnicos, salud humana privada, oficinas administrativas y de soporte y alojamiento/vivienda) y por empresas comerciales.

En el tercer nivel se presentan tres conglomerados: El primero conformado por empresas de manufactura, restaurantes y bares e investigadores privados y seguridad; el segundo, por agencias de viajes, educación superior privada, desarrollo y sistemas de información, telecomunicaciones, y profesionales, científicos y técnicos, y el tercero por empresas comerciales, alojamiento/vivienda, juegos de azar y actividades deportivas y recreativas.

En el cuarto nivel se presentan dos conglomerados: el primero formado por empresas de servicios de transmisión y producción de TV, actividades de publicación, agencias de viaje y educación superior privada. El segundo conglomerado conformado por empresas de alojamiento/vivienda, juegos de azar, actividades deportivas y recreativas y actividades inmobiliarias.

En el quinto nivel se presentan cinco conglomerados pequeños: el primero conformado por empresas manufactureras y restaurantes y bares; el segundo, por empresas de producción de películas, programación y transmisión de TV y actividades de publicación; el tercero, por agencias de viaje, educación superior privada y desarrollo y sistemas de información; el cuarto, por empresas de telecomunicaciones, profesionales, científicos y técnicos, salud humana privada y oficinas administrativas y de soporte, y el quinto y último grupo conformado por empresas de juegos de azar, actividades deportivas y recreativas, actividades inmobiliarias, e-mail y almacenamiento, y otros servicios y actividades. En la Tabla 6 se presentan los resultados de la agrupación:

Tabla 6. Historial de conglomeración

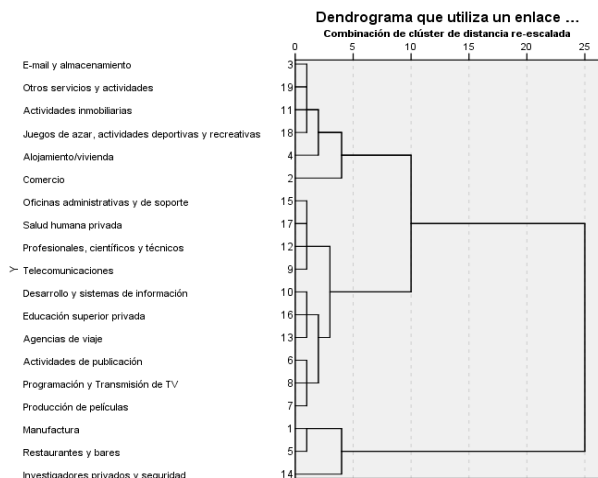
Etapa	Clúster combinado		Coeficientes	Primera aparición del clúster de etapa		Etapa siguiente
	Clúster 1	Clúster 2		Clúster 1	Clúster 2	
1	3	19	7,050	0	0	7
2	15	17	12,420	0	0	6
3	10	16	16,360	0	0	4
4	10	13	37,200	3	0	12
5	1	5	51,290	0	0	16
6	12	15	67,010	0	2	8
7	3	11	72,325	1	0	10
8	9	12	88,697	0	6	14
9	6	8	90,020	0	0	11
10	3	18	110,957	7	0	13
11	6	7	129,470	9	0	12
12	6	10	184,283	11	4	14
13	3	4	220,463	10	0	15
14	6	9	370,637	12	8	17
15	2	3	545,414	0	13	17
16	1	14	566,155	5	0	18
17	2	6	1384,509	15	14	18
18	1	2	3729,534	16	17	0

Fuente: Elaboración propia.

La Figura 4 presenta el dendrograma en el cual se pueden apreciar claramente las relaciones de agrupación entre los datos e incluso entre grupos de ellos, aunque no las relaciones de similitud o cercanía entre categorías.

VI. DISCUSIÓN

A partir de la clasificación, se identificaron algunos patrones de agrupamiento: el uso de las TIC en empresas manufactureras y de servicios difiere del uso que hacen las empresas comerciales. Los resultados permiten ver que, en el país, no todas las empresas comerciales cuentan con una página web y no todo el personal que trabaja en ellas usa computador e internet para realizar sus funciones. Las empresas de servicios de investigadores privados y seguridad son las que registran las tasas

**Fig. 4.** Dendrograma de TIC

Fuente: Elaboración propia.

más bajas de uso de computador e internet por parte de su personal.

Por otra parte, el personal que trabaja en empresas manufactureras es el que menos emplea computador e internet. Esta situación puede entenderse con claridad, porque la mayoría de actividades que se realizan en este tipo de empresas son actividades manuales (pegar, coser, tejer, bordar, troquelar, cortar, pulir, pintar, ensamblar, etc.) que no requieren su utilización.

VII. CONCLUSIONES

Gracias a la agrupación jerárquica es posible identificar objetos basados en las similitudes que presentan. En este trabajo, permitió agrupar los tipos de empresas colombianas (manufactureras, comerciales y de servicios) que tienen la mayor similitud en el uso de las TIC, específicamente en el uso del internet y de los computadores.

La agrupación jerárquica, a diferencia de otras técnicas de agrupamiento, tiene la ventaja de que puede utilizar cualquier tipo de distancia (euclídea, euclidiana tothesquare, Manhattan, Mahalanobis, máxima similitud, coseno), lo que permite realizar mejores agrupaciones.

El uso de las TIC en los procesos de producción facilita la labor del personal, reduce tiempos e incrementa la eficiencia. En efecto, las TIC se constituyen en una poderosa herramienta para acelerar el crecimiento económico de los países en desarrollo.

Las empresas deben ser flexibles y capacitar a todo el personal para adaptarse a los mercados cambiantes y variados. Deben estar presentes en las redes sociales: Twitter, LinkedIn, Facebook, Youtube e Instagram. Deben monitorear a la competencia actual y la potencial para que puedan enfrentar cambios. Deben planificar para innovar, dedicando los recursos necesarios, principalmente tiempo y personas. Analizar los comentarios de los clientes en las redes sociales, utilizar claves seguras, hacer backups de la información, cuidar la propiedad intelectual y capacitar a los empleados en seguridad informática.

REFERENCIAS

- [1] DANE, “Indicadores básicos de tenencia y uso de Tecnologías de la Información y Comunicación en empresas 2016”, 28 de diciembre de 2017. Disponible en: http://www.dane.gov.co/files/investigaciones/boletines/tic/bol_empresas_2016.pdf.
- [2] L. H. Gutiérrez and L. F. Gamboa, “Determinants of ICT Usage among Low-Income Groups in Colombia, Mexico, and Peru”, *Inf. Soc.*, vol. 26, N.º 5, pp. 346-363, Sep. 2010.
- [3] M. de Comunicaciones, “Plan Nacional de Tecnologías de la Información y las Comunicaciones”, 2008. Disponible en: <http://www.eduteka.org/pdfdir/colombiaplannacionaltic.pdf>
- [4] DANE, “Indicadores básicos de TIC en empresas (2016-2017)”, 2018. Disponible en: <https://www.dane.gov.co/index.php/estadisticas-por-tema/tecnologia-e-innovacion/tecnologias-de-la-informacion-y-las-comunicaciones-tic/indicadores-basicos-de-tic-en-empresas>
- [5] J. C. M. Cañizares, M. F. W. Díaz, G. H. Atondo, B. A. Sánchez, y H. Montiel, “Impacto de la gestión de los procesos colaborativos utilizando las TIC en empresas de manufactura”, *Int. Rev. Bus. Res. Pap.*, vol. 8, N.º 3, 2012.
- [6] T. DeStefano, R. Kneller, and J. Timmis, “Broadband infrastructure, ICT use and firm performance: Evidence for UK firms”, *J. Econ. Behav. Organ.*, vol. 155, pp. 110-139, 2018.
- [7] J. Altés, “Papel de las tecnologías de la información y la comunicación en la medicina actual,” *Semin. la Fund. Española Reumatol.*, vol. 14, N.º. 2, pp. 31-35, 2013.
- [8] B. B. Hughes, D. Bohl, M. Irfan, E. Margolese-Malin, and J. R. Solórzano, “ICT/Cyber benefits and costs: Reconciling competing perspectives on the current and future balance”, *Technol. Forecast. Soc. Change*, vol. 115, pp. 117-130, 2017.
- [9] M. Alderete, “TIC y productividad en las industrias de servicios en Colombia”, *Lect. Econ.*, vol. 5, N.º 77, pp. 163-188, 2012.
- [10] D. Oliveros Contreras and G. M. Martínez, “Efecto de las TIC sobre la gestión de las empresas hoteleras afiliadas a Cotelco de Bucaramanga (Santander, Colombia)”, *Rev. EAN*, N.º 83, pp. 15-30, 2017.

- [11] L. Paulise, “¿Cuál es el impacto de las TIC en las pequeñas empresas?”, 2019. Disponible en: <https://destinonegocio.com/co/emprendimiento-co/impacto-tic-pequenas-empresas/>.
- [12] M. Zapata, “Con la aplicación en TIC tendremos en Colombia empresas más competitivas”, 2019. Disponible en: <https://www.edatel.com.co/empresas/blog-empresas/414-con-la-aplicacion-en-tic-tendremos-en-colombia-empresas-mas-competitivas>
- [13] D. Grant and B. Yeo, “A global perspective on tech investment, financing, and ICT on manufacturing and service industry performance”, *Int. J. Inf. Manage.*, vol. 43, pp. 130-145, 2018.
- [14] ITU World Telecommunication, “Global ICT developments 2001-2018”, 2019. Disponible en: <https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx>.
- [15] E. Herman, “The importance of the manufacturing sector in the Romanian economy”, *Procedia Technol.*, vol. 22, pp. 976-983, 2016.
- [16] PROCOLOMBIA, “La Manufactura en Colombia”, 2018. Disponible en: <http://www.procolombia.co/compradores/es/explore-oportunidades/manufactura-en-colombia>.
- [17] Invest Pacific, “Invest Pacific, after more european investment for the Manufacturing Sector”, 2 de mayo de 2019. Disponible en: <http://www.investpacific.org/en/pressroom.php?id=828>.
- [18] J. E. Giraldo, “El comercio en la economía de Colombia”, 2017. Disponible en: <https://www.gestiopolis.com/el-comercio-en-la-economia-de-colombia/>.
- [19] J. Sáenz, “Economía crece 2,7 % en 2018 por sector público y comercio”, 2019. Disponible en: <https://www.elespectador.com/economia/economia-crece-27-en-2018-por-sector-publico-y-comercio-articulo-842438>.
- [20] I. U. Din, M. Guizani, J. J. P. C. Rodrigues, S. Hassan, and V. V. Korotaev, “Machine learning in the Internet of Things: Designed techniques for smart cities”, *Futur. Gener. Comput. Syst.*, 2019.
- [21] N. Peek, C. Combi, R. Marin, and R. Bellazzi, “Artificial Intelligence in Medicine. Thirty years of artificial intelligence in medicine (AIME) conferences: A review of research themes”, *Artif. Intell. Med.*, vol. 65, N.º 1, pp. 61-73, 2015.
- [22] T. O. Ayodele, “Machine learning overview”, in *New Advances in Machine Learning*, IntechOpen, 2010. Disponible en: <https://www.intechopen.com/books/new-advances-in-machine-learning/machine-learning-overview>
- [23] S. B. Kotsiantis, I. Zaharakis, and P. Pintelas, “Supervised machine learning: A review of classification techniques”, *Emerg. Artif. Intell. Appl. Comput. Eng.*, vol. 160, pp. 3-24, 2007.
- [24] S. A. Mondal, “An improved approximation algorithm for hierarchical clustering”, *Pattern Recognit. Lett.*, vol. 104, pp. 23-28, 2018.
- [25] T. Bock, “What is Hierarchical Clustering?”, 2019. Disponible en: <https://www.displayr.com/what-is-hierarchical-clustering/>
- [26] V. Gil and I. Lopera, “Classifying death causes with hierarchical clustering: The Colombian Case”, *Middle-East J. Sci. Res.*, vol. 26, N.º 4, pp. 438-443, 2018.
- [27] T. Samarasinghe, T. Gunawardena, P. Mendis, M. Sofi, and L. Aye, “Dependency Structure Matrix and Hierarchical Clustering based algorithm for optimum module identification in MEP systems”, *Autom. Constr.*, vol. 104, pp. 153-178, 2019.
- [28] I. A. Pagnuco, J. I. Pastore, G. Abras, M. Brun, and V. L. Ballarin, “Analysis of genetic association using hierarchical clustering and cluster validation indices”, *Genomics*, vol. 109, N.º 5, pp. 438-445, 2017.

Capítulo VII

Obtención del Conocimiento para Análisis de Requisitos en el Desarrollo de Software

Diana María Montoya Quintero - dianamontoya@itm.edu.co

Sonia Jackeline Moreno - soniamoreno@itm.edu.co

Instituto Tecnológico Metropolitano ITM, Medellín, Colombia

I. INTRODUCCIÓN

Las Tecnologías de Información y Comunicación (TIC) y la Gestión del Conocimiento (GC) hacen parte de las herramientas de creatividad que se pueden aplicar al desarrollo de software en su ciclo de vida de desarrollo. Ambas herramientas han presentado una serie de iniciativas que han indagado problemáticas y necesidades específicas para dar respuesta a los requerimientos funcionales de un producto de software en una de sus primeras etapas, como es el análisis de requisitos. Es de entender que estas iniciativas o acciones varían de acuerdo al entorno y el contexto, y permiten transformaciones al ser implementadas. En este artículo se propone un paso a paso para la aplicación de herramientas que gestionan el conocimiento en productos empresariales de software.

Los términos de Herramientas Creativas, TIC, informática tecnológica (IT) y GC han venido tomando relevancia en los últimos años, en especial porque son estrategias que han puesto el foco en algunas organizaciones. Estos conceptos, que aún están en construcción en el mundo académico como se observará más adelante, tienen varios términos asociados, y es importante tener un acercamiento a estos conceptos, los cuales son los principales referentes en esta investigación. En este artículo también se exponen aquellas estrategias o acciones que se circunscriben al contexto organizacional.

A. Herramientas Creativas

De acuerdo con lo planteado en [1], “las capacidades de infraestructura tecnológicas son herramientas mediadoras en el establecimiento

de las políticas y estrategias empresariales, lo cual permite una mejor gestión del capital humano”. Hay que mencionar además que no todas las herramientas deben estar contempladas o interactuadas por internet. Entonces, la palabra *herramientas* hace referencia a:

- Objeto que se utiliza para trabajar en oficios o realizar un trabajo manual.
- Conjunto de instrumentos [2].

Y *creatividad* hace referencia a:

- Facultad de crear: esta actividad desarrolla la creatividad del niño [2].
- Capacidad de creación.
- La creatividad, denominada también pensamiento original, pensamiento creativo, inventiva, imaginación constructiva o pensamiento divergente, es la capacidad de crear, innovar, generar nuevas ideas o conceptos o nuevas asociaciones entre ideas y conceptos conocidos, que normalmente llevan a conclusiones nuevas, resuelven problemas y producen soluciones originales y valiosas [3].

Con respecto a los términos de herramientas y creatividad, las herramientas creativas, como su nombre lo indica, son instrumentos usados con el fin de crear, aunque no es pericia de todo ser humano, conviene subrayar que la creatividad parte de ideas que surgen de una necesidad o condición de acuerdo con el entorno en que se presente, si usamos herramientas que permitan el desarrollo de esas nuevas ideas como estímulo al pensamiento, para lograr mejoras o impulsar estrategias en cualquier organización.

B. Las Tecnologías de Información y Comunicación (TIC)

Las TIC “son el conjunto de tecnologías que permiten el acceso, producción, tratamiento y comunicación de información presentada en diferentes códigos (texto, imagen, sonido, etc.)” [4]. La Unesco [5] afirma que las TIC tienen un “rol fundamental en el acceso universal a la educación, la igualdad en la instrucción, la enseñanza y el aprendizaje de calidad”. Las TIC han intervenido de manera significativa el sistema educativo creando nuevas tendencias pedagógicas, una de ellas es el aprendizaje híbrido (mixto, combinado o *blended learning*) que combina diferentes formatos de tecnologías web con diferentes enfoques pedagógicos y facilita promover en el estudiante o empleado la cultura de aprendizaje autónomo (aprender a aprender solo, en equipo o combinado), que se traduce en trabajo independiente por fuera del aula, el espacio de trabajo o la organización [6]. Por otra parte, en [7] se plantea que las TIC contribuyen a la construcción de conocimiento en el sentido que existe una gran facilidad para acceder a más información, porque se incrementan las posibilidades de interactividad y conectividad, se utilizan formatos multimedia, lo que redundará en la generación de contenidos.

C. Tecnologías Informáticas

Por tanto, la tecnología es un conjunto de técnicas y procedimientos que se diseñan para solucionar un determinado problema [8]. La informática, por su parte, es conocida como una ciencia y está ligada íntimamente con las computadoras. En relación con lo anterior, las tecnologías informáticas (TI) son el estudio, diseño, desarrollo, innovación puesta en práctica, ayuda o gerencia de los sistemas informáticos computarizados, particularmente usos del software, en general, del uso de computadoras y del software electrónico. Asimismo, son la posibilidad de convertir, almacenar, proteger, procesar, transmitir y recuperar la información [9]. La tecnología cumple, entonces, un papel

importante en todas las áreas del quehacer humano, por lo que se ha visto la necesidad de introducirla en el área de la educación, para poder mejorar el proceso de enseñanza-aprendizaje [10].

D. Gestión del Conocimiento

En algunas investigaciones se [11] expresa la importancia que ha alcanzado el análisis de la gestión del conocimiento (GC) en las organizaciones de todo tipo. La GC se fundamenta en la posibilidad de generar modelos que permitan la implementación de tecnologías que facilitan la gestión exitosa de los planes, tanto estratégicos como operativos, mediante transacciones de información y conocimiento.

De acuerdo con el acercamiento a los conceptos, se puede decir que existe una estrecha relación entre ellos, lo cual permite crear metodologías, agregar valor, generar ideas, solucionar problemas, innovar en la organización, etc. La importancia de la herramienta creativa mediada por las TIC en la GC, radica en que existen en el mundo web cantidades de herramientas y estrategias que facilitan la GC en cualquier organización, sin embargo, algunas organizaciones no hacen uso de ellas porque cuando se esbozan no se evidencia un paso a paso de su metodología o uso. En este capítulo se pretende plantear herramientas creativas para la GC.

II. MATERIALES Y MÉTODOS

Con el fin de proponer un paso a paso para la aplicación de herramientas para gestionar el conocimiento en productos empresariales de software, se planteó una metodología de tipo descriptivo mediante una recolección de información sobre conceptos y variables sin pretender hallar una correlación entre ellos [12]. La investigación también es de carácter exploratorio, porque se recolectan datos en fuentes secundarias. Otro aspecto importante que se debe resaltar es el enfoque, que se inscribe más en un método cualitativo porque sólo se trataron datos no numéricos.

Con referencia a lo mencionado, se tuvieron en cuenta fuentes secundarias tales como bases de datos científicas y especializadas, libros, artículos científicos e informes técnicos, además de páginas web. Con una detallada revisión de la información obtenida de la literatura, se analizaron casos similares, antecedentes y perspectivas acerca del tema, y de acuerdo con lo recolectado se delimitó el tema de estudio y se le dio forma a la investigación.

A este efecto, la investigación se desarrolló en dos fases: En la fase 1 constituye un componente metodológico, el paso a paso de implementación, y en la fase 2 se exponen las herramientas, materiales y diseño.

III. RESULTADOS

A. Fase I: Metodología de implementación de herramientas

Cada que se lleva a cabo un proceso de implementación de herramientas mediadas con TIC, se requiere una fase pedagógica de socialización en el área que va a ser impactada por aquellas que la organización decida elegir para gestionar su conocimiento. Es importante concertar discursos que hagan parte de este ejercicio de comunicación desde diversos ámbitos de la organización: administrativos, económicos, ambientales, técnicos, etc.; esta es la razón de ser de esta metodología, pues es preciso dar a conocer un mensaje claro y concreto de lo que comprende y representa una herramienta mediada con TIC. Los cinco pasos de la metodología se observan en la Figura 1.

Valoración: Es fundamental identificar la necesidad, el conflicto o el problema que experimentan los productos empresariales de software, si este reconocimiento lo hace la organización, debe existir una solicitud por parte del área o proceso encargados, porque son ellos quienes conocen el vacío en el puesto de trabajo o estudio.



Fig. 1. Pasos de la metodología

La retroalimentación de trabajo del equipo encargado del área o proceso que se van a intervenir, debe consolidar la información de diagnóstico la cual permite iniciar el proceso de formulación.

La valoración se realiza conjuntamente entre el personal administrativo y los encargados, por tanto se busca establecer problemáticas y esbozar posibles soluciones por medio de un mapeo de las necesidades. Este mapeo permite identificar y reconocer las particularidades de las organizaciones, y consiste en recorrer la organización para registrar las diferentes lecturas del contexto, donde se hace una caracterización de quiénes conforman la organización, se analizan sus espacios de recreación y equipamientos, los usos, su infraestructura, sus relaciones sociales y laborales, etc.

Formulación: Se realiza un análisis profundo de las necesidades, problemas y/o falta de recursos que manifiestan las personas de la organización y las que integran el área o proceso que se va a intervenir, en los espacios de participación que se disponen como: reuniones, recorridos en la organización, entrevistas, etc. En este punto, los aportes, el conocimiento y

la experiencia compartida sirven como insumo fundamental para definir y precisar la propuesta de selección de herramientas creativas que se han de usar para gestionar su conocimiento, priorizando los problemas y necesidades identificadas bien sea de manera individual o combinadas. En este paso, dentro del marco de selección de herramientas y los esquemas de actuación que buscan resolver las necesidades del área de la organización identificada en el paso previo por los empleados y/o estudiantes y el equipo interdisciplinar, todo está sujeto a la validación, teniendo en cuenta las diferentes recomendaciones, observaciones y demás información recopilada, lo cual consiste en verificar una serie de premisas que permitan legitimar las acciones que se van a realizar y que sean fáciles de comprender.

Implementación: Es el paso operativo del proyecto que busca dar inicio a la propuesta que se concretó desde los distintos actores. Además de verificar la eficiencia y efectividad de la implementación de la herramienta que la organización decida seleccionar, se hace un registro de las limitaciones y beneficios de la herramienta creativa, con el fin de identificar oportunamente las fortalezas y debilidades de los procesos de ejecución e interactividad, con el objetivo de realizar modificaciones para lograr una adecuada gestión.

Seguimiento: Se presenta a la organización los resultados logrados, en este paso se generan y analizan opiniones o sugerencias acerca de la herramienta que se puso en práctica como estrategia para conservar y crear conocimiento; asimismo, el seguimiento permite retroalimentar lo propuesto para mejorarlo o replicarlo en otras áreas de la organización. En el proceso de evaluación, posibilita establecer, a la luz de los objetivos, la relevancia, eficiencia, eficacia e impacto de la herramienta. De esta manera, la evaluación va más allá del simple monitoreo, ya que asume que el plan de acción hace parte de una hipótesis con respecto a la ruta que se debe seguir para lograr los objetivos. La socialización y presentación de los resultados a la organización incluye un análisis conjunto, y en

este espacio se reconocen las diferentes sugerencias sobre las estrategias y/o herramientas que se aplicaron. Es importante realizar esta socialización para evidenciar ante la organización los aspectos que se fortalecieron y permitir una retroalimentación que redunde en el éxito de la propuesta.

Componente comunicacional: En este paso es muy importante crear un plan de comunicaciones que difunda oficialmente la información de los resultados de la estrategia implementada a toda la organización. Es el soporte para que se tengan las herramientas conceptuales a fin de poder socializar las experiencias y para dar respuesta a las demás áreas o procesos, medios, empleados o estudiantes que busquen información sobre el tema (ver Figura 2).



Fig. 2. Acciones de comunicación

Dentro del plan de comunicación se deben prever las inquietudes que seguramente consultará la organización a los encargados, y parte de este plan consiste en tener material informativo que les dé respuesta.

B. Fase 2: Herramientas, materiales y diseño

En la segunda fase, las herramientas que se describen a continuación forman parte de lúdicas y experiencias establecidas o diseñadas para crear y conservar el conocimiento.

Metodología LSP:

Es en el juego y solo en el juego que el niño y el adulto como individuos son capaces de ser creativos y usar el total de su personalidad, y solo al ser creativo, el individuo se descubre a sí mismo.

Donald Woods [13].

Lego Serious Play (LSP) es una metodología de comunicación, trabajo en equipo y resolución de problemas utilizada en más de 29 países para el desarrollo de las organizaciones. A través de las preguntas de un facilitador, cada participante construye su propio modelo tridimensional utilizando piezas especializadas de Lego, y además construye estrategias para la organización como se observa en la Foto 1.



Foto 1. Estrategia

Fuente: tomada por los autores en una experiencia de aula

Todos construyen, todos comparten, todos reflexionan, todos hablan desde el modelo.

Pasos:

- Pregunta o reto
- Construcción
- Metáfora
- Reflexión

Habilidad para construir, ejemplo:

- 20 bloques
- 2 minutos
- Construir el auto del futuro
- ¿Qué necesita una persona para ser creativa e innovadora?
- ¿Y cómo será nuestro modelo compartido?
- ¿Cuáles son los agentes que pueden favorecer nuestro modelo?
- ¿Cómo se conectan?
- ¿Qué aprendimos? [14]

En la Foto 2 se observa una experiencia en aula. Las fotografías 1 y 2 fueron tomadas en una experiencia de aula donde se logró una capacitación a estudiantes de maestría, empresas y docentes.



Foto 2. Trabajo colaborativo a partir de un reto

Fuente: tomada por los autores en una experiencia de aula

Los seis sombreros para pensar: La técnica de los seis sombreros para pensar (Foto 3), desarrollada por Edward de Bono, es una herramienta de comunicación y razonamiento muy efectiva [15]. Cuando una persona se enfrenta a un proceso de toma de decisiones, su mente comienza a contemplar una serie de razones y emociones que la llevan a optar por una alternativa. Ese extraño proceso mental en que el individuo coteja las ventajas, los inconvenientes, los hechos, los sentimientos y otra serie de informaciones relevantes, resulta muy difícil de explicar o de simular. En consecuencia, lo que se da de forma natural en el cerebro humano suele resultar extremadamente complejo cuando, por ejemplo, tratamos de aplicarlo a una decisión colectiva.



Foto 3. Seis sombreros para pensar [16]

Pasos:

- Inicia con el sombrero blanco, el color de la neutralidad, con el cual entran en juego solo los datos, hechos y cifras comprobables.
- Luego viene el sombrero de color verde, en el que se define la lista de hechos que se están presentando en el proceso, qué información se está omitiendo y cómo se puede acceder a ella.
- Posteriormente, sigue con el sombrero rojo, la idea es que todos expresen sus sentimientos hacia los hechos que se identificaron en la primera etapa, cuál es su sentir, qué le gusta, qué le disgusta, etc.
- El siguiente es el sombrero de color negro, con él hay que ser pesimista, cuáles son las desventajas y los riesgos que traen los hechos identificados con el sombrero blanco.
- Luego utilice el sombrero amarillo, este actúa de manera contraria al sombrero negro; sea optimista con la situación, cuáles son las ventajas y los beneficios que se obtendrán si se mejoran los hechos que no permiten que las cosas funcionen.
- Siga con el sombrero verde, que implica focalizarse en las alternativas; existen nuevas opciones para mejorar la situación que se presenta, generar nuevas ideas para solucionar los problemas.
- Por último, utilice el sombrero azul, este nos lleva a focalizarnos en el control y el proceso, se usa para pensar en pensar, acá se hace un resumen de lo sucedido, se sacan conclusiones y se toman decisiones

La técnica propuesta ha sido replicada por muchas organizaciones, lo más importante es que los participantes logren apropiarse del papel que concierne, así no solo se dará cuenta de todo el conocimiento que posee, sino también de todas las ideas que pueden resultar del ejercicio, además, ayudar a fortalecer los procesos de la organización. En [17] se puede ver claramente cuáles son las estrategias de cada sombrero con ejemplos.

Objetos Virtuales de Aprendizaje: La herramienta Exelearning es una de las herramientas para la construcción de Objetos Virtuales de Aprendizaje OVA. Un OVA es un recurso digital que puede ser reutilizado en diferentes contextos. Pueden ser cursos, cuadros, fotografías, películas, vídeos y documentos que posean claros objetivos educativos [18]. En la Figura 4 se observan las características de la interfaz de un OVA (adaptado con Exelearning), al lado izquierdo se encuentra el menú de un OVA de matemáticas básicas, por ello se observan los temas correspondientes, cada uno de los títulos se despliega con su respectiva definición y actividades propuestas, tales como talleres, videos, autoevaluaciones, PDF y diapositivas.



Fig. 3. El ejemplo de la interfaz de un OVA

Pasos del diseño:

- Selección de la información (conocimientos previos), definición del tema, experiencia, capacitación o aprendizaje que la organización desea preservar en el tiempo.
- Selección o construcción de videos de tipo monoconceptual.
- Selección de autoevaluaciones, en la plantilla se construye: Falso-cierto, el ahorcado, selección múltiple.
- Construcción de teoría en PDF y diapositivas.

Una vez se cuente con toda la información se procede a la construcción del OVA con la herramienta eXeLearning [19], que se puede descargar de forma gratuita, y se encuentra todo lo relacionado para su

construcción [19]. Se debe agregar que cuando se diseña un OVA, se debe crear una licencia Creative Commons que permita la divulgación libre de la información plasmada por la organización.

Objetos interactivos de aprendizaje (OIA): Descartes JS es una de las herramientas para la construcción de los OIA, que permite contener información, videos, PDF, diapositivas, autoevaluaciones, etc. Entre los elementos que contiene un OIA se encuentran:

- *Interactividad:* Moreno y Mayer, citados en [20], proponen cinco tipos de interactividad: diálogo, control, manipulación, búsqueda y navegación. Estos a su vez están basados en cinco principios: actividades guiadas, reflexión, realimentación, control y preentrenamiento o experiencia. Es decir, la interactividad es la característica que permite al usuario establecer un canal de doble vía para recibir información de parte del OIA y entregar nueva información como respuesta al mismo OIA.
- *Aleatoriedad:* Ferney y Velásquez [21] plantean que “la idea de aleatoriedad se introduce preferentemente de un modo descriptivo en la educación, resultando fundamental el papel del lenguaje utilizado para comunicar resultados”, más aún, es una característica que permite al usuario variar algún parámetro bien sea cuantitativo o cualitativo en el problema o ejercicio, el tipo de ejercicio o el tipo de pregunta, de manera que se planteen situaciones distintas cada vez.
- *Retroalimentación inmediata:* Canabal y Margalef [22] expresan que “en un contexto de evaluación como aprendizaje desarrollados”, la retroalimentación también incluye la aparición de mensajes con pistas para intentar nuevamente la solución o una explicación de cuál sería la solución según las condiciones planteadas en la escena del OIA.

- *Portabilidad:* Es la posibilidad de acceder a los recursos desde diferentes dispositivos y sistemas operativos (teléfonos móviles, tabletas y portátiles).

En la Figura 4 se observan los pasos para el diseño de los OIA y en la Figura 5 se observa un ejemplo de la interfaz de un OIA en forma de libro digital.



Fig. 4. Pasos del diseño OIA



Fig. 5. Interfaz OIA (Tomado de: <https://gruponomon.itm.edu.co/pages/teoria.html>)

En la página del proyecto Descartes están contenidas las plantillas para que los interesados en el diseño de OIA puedan acceder a la información y realizar el diseño que consideren pertinente para el desarrollo de su objetivo.

Las evidencias de los resultados obtenidos de las herramientas propuestas, se basa en la información de las experiencias de aula. Cada una de las herramientas se implementó en el Instituto Tecnológico Metropolitano (ITM) en el diplomado de Didáctica Universitaria, el cual a la fecha tenía cuatro grupos. A estos después de vivir su experiencia con las herramientas, se les solicitó responder una encuesta de satisfacción, y con ellos

mismos se socializó la experiencia. Como respuesta se obtuvieron los siguientes resultados de acuerdo a cada experiencia.

Metodología LSP: Las experiencias de aula propiciaron espacios de aprendizaje, transferencia de conocimiento, confrontación de conocimiento, comunicación activa, entre otros. Algunos aportes de los participantes fueron los siguientes:

- El recrear escenarios de contextos reales de una organización donde se permita el aprendizaje a través del juego, generó experiencia que puede permanecer más fácil en uno.
- Cuando se involucró la experiencia didáctica en el aula se logró evidenciar un aprendizaje más significativo acerca del uso y aplicación de la metodología, partiendo de un problema real de una organización.

Los seis sombreros para pensar: En la aplicación de esta herramienta se desarrolló la lúdica llamada “Aeroplano de papel”, muy utilizada en otras lúdicas de aprendizaje [23], ya que si cada individuo es capaz de inventar o improvisar una aventura lúdica original, ésta se apoya en los cimientos de la evolución de todo lo que ha venido generando el colectivo humano al que pertenece.

En la Foto 4 se observa uno de los grupos en el cual ya se realizaban los análisis a partir de la herramienta seis sombreros para pensar.



Foto 4. Experiencia de Aula

Al cambiar los roles de la organización que representaban en la empresa, concluyeron que los ingenieros y directivos rara vez se apropian del trabajo del operario; por tal motivo, no se logran evidenciar claramente los cuellos de botella en el área o proceso.

Objetos Virtuales de Aprendizaje (OVA): En la Tabla 1 uno se muestra la encuesta aplicada a uno de los grupos con sus respectivos porcentajes y algunas apreciaciones referentes a los OVA como medio de aprendizaje.

Tabla 1. Encuesta sobre los OVA

Pregunta	Porcentaje de resultados		
	Sí	No	Otra
1. ¿La temática del OVA es clara?	95,9 %	4,1 %	0
2. ¿Considera que con este medio de divulgación se gestiona el conocimiento?	100 %	0	0
3. ¿La lectura en el OVA le permite tener un concepto cultural sobre el tema que se pretende aprender?	100 %	0	0
4. ¿Ve este medio digital como una oportunidad de cambio de acuerdo con el tradicional?	95,9 %	4,1 %	0
	Solo medios digitales	La presencia de un docente	La combinación de las dos anteriores (medios digitales y el docente)
5. ¿Qué prefiere para su aprendizaje?	0	2 %	98 %

Fuente: Elaboración propia.

La justificación de la respuesta 5 por parte de los involucrados se describe a continuación:

- Los medios digitales son una herramienta, la cual nosotros como estudiantes tenemos más acceso a diferentes puntos de vista del mismo tema, además la presencia del docente y su explicación hace que teóricamente se tengan mejores bases a la hora de buscar apoyos visuales y temas en la red.
- Siempre necesitamos la ayuda de un profesor
- La forma de aprendizaje presencial es la manera por la cual aprendo e interactué más con la materia, pero ahora los medios digitales son una gran manera de aprender y otra alternativa para cambiar de ambiente.
- Considero que tener un docente para el aprendizaje es vital, ya que este tiene un papel muy importante tanto con las horas de clases, como con la ayuda que nos presta para solucionar dudas. Los medios digitales nos pueden ayudar con otras opiniones y formas de enseñanza para enriquecer nuestro conocimiento.
- Porque es útil ya que nos sirve para estudiar en la universidad como en la casa.
- El docente guía el conocimiento y el trabajo independiente y asequible, son una combinación primordial para el aprendizaje.

Objetos Interactivos de Aprendizaje (OIA):

Tabla 2. Encuesta sobre los OIA

Pregunta	Porcentaje de respuestas					
	Muy clara	Clara	Algunos aspectos son claros	La temática es poco clara	No se entiende la temática	Muy atractiva
1. ¿La temática del OIA es clara?	38 %	60 %	0	0	0	2 %
	Nada atractivo para GC	Poco atractivo para GC	Atractivo para GC	Es un medio bastante atractivo para GC	Excelente para la GC	Atractiva
2. ¿Considera que con este medio de divulgación se gestiona el conocimiento?	0	4 %	12 %	60 %	22 %	2 %
3. ¿La lectura en el OIA le permite tener un concepto cultural sobre el tema que se pretende aprender?	0	0	18 %	42 %	38 %	2 %
	No creo que cambie la educación tradicional	En algunos casos puede hacer el cambio	Es una oportunidad atractiva para el cambio	Le veo muchas oportunidades para el cambio	Es la mejor oportunidad para cambiar	Muy atractiva
4. ¿Ve este medio digital como una oportunidad de cambio en la educación tradicional?	0	10 %	20 %	44 %	24 %	2 %
5. ¿Qué prefiere para su aprendizaje?	100 %					

Fuente: Elaboración propia.

La justificación de la respuesta 5 por parte de los involucrados se describe a continuación:

- Considero que las herramientas virtuales son una muy buena opción en la actualidad, sin embargo, es importante la presencialidad con el fin de afianzar conocimientos y aprendizajes.
- Los medios digitales son herramientas para el aprendizaje, pero la formación en criterio siempre debe estar dirigida por un docente.
- Los medios interactivos y digitales permiten acercar el conocimiento como un aire más calmado y con menos intimidación que un libro tradicional. Sin embargo, el acompañamiento del docente es indispensable para aclarar dudas y ambigüedades, ya que a un texto no se le puede pedir aclaración o explicación.
- Por el medio digital se pueden adquirir varios conocimientos, familiarizarse con los temas, y con la ayuda del profesor aprender y resolver dudas e inquietudes y ayudar a comprender mejor los temas.
- Este método me parece muy interesante, porque podemos leer el concepto, ver la información precisa del tema y realizar ejercicios, lo cual permite saber si se está aprendiendo.
- El docente siempre va ser necesario, porque es la mejor forma de transmitir el conocimiento, además, cuando se está en clase se pueden solucionar dudas con mayor claridad, pero los medios digitales son un gran apoyo para el estudio independiente.
- Prefiero la combinación de ambas, ya que si solo fuera lo digital y surgieran dudas no tendríamos ayuda que el docente puede ofrecer, la combinación de las dos ayuda a reforzar conocimiento.

IV. CONCLUSIONES

La pretensión de los autores con estas lúdicas de obtención de conocimiento es encontrar soluciones reales que permitan llevar las necesidades a entornos tecnológicos, para ser codificados por procesos sistemáticos en el desarrollo del software. Se reconoce que aún es un cuello de botella obtener

los requerimientos del usuario de una forma práctica y divertida que permita el diseño y desarrollo del producto de forma dinámica y más humana.

Una vez se obtiene la información, esta se debe convertir en conocimiento y una vez se obtiene el conocimiento este se transfiere para generar un nuevo conocimiento que, muchas veces, en el área de software es más práctico, preciso y exacto en el momento de ser consultado por cualquier usuario interesado.

Por medio de la investigación realizada se planteó un paso a paso para la implementación de las TIC y la GC, sin embargo, en el mundo académico y empresarial existen muchas más herramientas que permiten conservar el conocimiento.

RECOMENDACIONES

Al utilizar la aplicación multimedia interactiva como los OVA y los OIA, se obtiene una herramienta pedagógica interactiva y participativa para mejorar el aprendizaje de los interesados.

REFERENCIAS

- [1] L. E. Valdez Juárez, D. García Pérez, G. Maldonado, "TIC y la gestión del conocimiento como elementos determinantes del crecimiento de la Pyme", *Investig. Cienc.*, vol. 25, N.º 12, pp. 50-62, 2017.
- [2] WordReference, "Herramientas", Definición - WordReference.com, 2019. [En línea]. Disponible en: <https://www.wordreference.com/definicion/herramientas>. [Accedido en: 06-May-2019].
- [3] Significados.com, "Creatividad", 2017. [En línea]. Disponible en: <https://www.significados.com/creatividad/>. [Accessed: 06-May-2019].
- [4] C. Beloch Ortí, "Las Tecnologías de la Información y Comunicación (T.I.C.) en el aprendizaje", Universidad de Valencia, 2008.
- [5] Unesco, "Las TIC en la Educación", Organización de las Naciones Unidas para la

- Educación, la Ciencia y la Cultura, Unesco, 2017. [En línea]. Disponible en: <http://www.unesco.org/new/es/havana/areas-of-action/education/tic-en-la-educacion/>. [Accedido: 06-May-2019].
- [6] J. J. García, “Análisis del rendimiento académico en el estudio de los límites de funciones de variable real con el apoyo de Objetos Interactivos de Aprendizaje (OIA)”, 2017.
- [7] O. Chaves, L. Chaves y D. Rojas, “La Realidad del uso de las TIC y su mediación pedagógica para enriquecer las clases de inglés”, *Rev. Ensayos Pedagógicos*, vol. 10, N.º 1, pp. 159-183, 2015.
- [8] Tecnología & Informática, “¿Qué es la tecnología? ¿Qué es la tecnología en informática?”, 2019. [En línea]. Disponible en: <https://tecnologia-informatica.com/que-es-tecnologia-informatica/> [Accedido en: 08-May-2019].
- [9] EcuRed, “Tecnología Informática”, 2019. [En línea]. Disponible en: https://www.ecured.cu/Tecnología_Informática. [Accedido en: 08-May-2019].
- [10] D. J. Morales Castro y J. D. Menéndez Sarmiento, “La Web 2.0 en la enseñanza de los recursos educativos abiertos (REA). Diseño de una guía didáctica integradora”, Tesis de grado, Universidad de Guayaquil Facultad de Filosofía, Letras y Ciencias de la Educación, Guayaquil, Ecuador, 2018.
- [11] M. A. Zabaleta, M. I. Brito, L. E. Garzón, “Modelo de gestión del conocimiento en el área de TIC para una universidad del Caribe colombiano”, *Rev. Lasallista Investig.*, vol. 13, N.º 16, pp. 1-16, 2016.
- [12] R. Hernández, C. Fernández y M. del P. Baptista, *Metodología de la investigación*. México: McGraw Hill, 2010.
- [13] D. Woods, *Realidad y juego*. Barcelona: Gedisa, 1979.
- [14] Máximo Impacto S.A.S., “Máximo Impacto – Consultoría en estrategia de alto nivel”, 2018. [En línea]. Disponible en: <http://conmaximoimpacto.com/> [Accedido en: 06-May-2019].
- [15] E. De Bono, *Seis sombreros para pensar*. Barcelona: Granica, 1988.
- [16] M. Saavedra Seoane, “Los seis sombreros para pensar | designthinking.gal,” 2018. [En línea]. Disponible en: <https://designthinking.gal/tecnicas-de-creatividad-los-seis-sombreros-para-pensar/>. [Accedido en: 06-May-2019].
- [17] FormaLider Academy, “Seis sombreros para pensar - Edward de Bono - Resumen del libro en español”, 30-May-2016. [Video] Disponible en: <https://www.youtube.com/watch?v=Mq7mYdrUOqs>
- [18] Ministerio de Educación Nacional de Colombia (MEN), “Objetos de aprendizaje virtual”, MEN, 2015. [En línea]. Disponible en: <https://www.mineducacion.gov.co/1621/articulo-82739.html> [Accedido en: 06-May-2019].
- [19] A. Monje, “Tutorial, manual de eXeLearning”, 2016. [En línea]. Disponible en: http://exelearning.net/html_manual/exe_es/index.html [Accedido en: 06-May-2019].
- [20] I. D. Claros y R. Cobos, “Del vídeo educativo a objetos de aprendizaje multimedia interactivos: un entorno de aprendizaje colaborativo basado en redes sociales”, *Tendencias Pedagógicas*, vol. 0, N.º 22, pp. 59-72, 2013.
- [21] E. F. Montoya Velásquez, “Azar, aleatoriedad y probabilidad: Significados personales en estudiantes de educación media”, tesis de grado, CINDE, Universidad de Antioquia, 2011.
- [22] C. Canabal y L. Margalef, “La retroalimentación: La clave para una evaluación orientada al aprendizaje”, *Profesorado. Revista de Currículum y Formación de Profesorado*, vol. 21, N.º 23, pp. 149-170, 2017.
- [23] J. Herrador, “Los juegos tradicionales en la filatelia: Estudio praxiológico y multicultural de la actividad lúdica. Traditional Games in Stamps: Praxiological Analysis and Multicultural Study of Leisure Activities”, *Acción Motriz*, N.º 6, pp. 58-75, 2011.

Capítulo VIII

Analítica de Datos para Determinar la Afinidad Académica y Personal en Programas de Ingeniería del Politécnico Colombiano Jaime Isaza Cadavid

Mónica María Rojas - mmrojas@elpoli.edu.co

Alejandro Pelaez Piedrahita - alejandro_pelaez82101@elpoli.edu.co

Paula Andrea Vargas Alvarez - paula_vargas82111@elpoli.edu.co

Politécnico Colombiano Jaime Isaza Cadavid, Medellín, Colombia

I. INTRODUCCIÓN

En Colombia, se viene hablando con preocupación de las cifras de deserción escolar y de la necesidad de disminuirla. Se entiende por deserción escolar “el abandono del sistema educativo por parte de los estudiantes, provocado por una combinación de factores que se generan, tanto en la escuela como en contextos de tipo social, familiar e individual” [1].

La deserción en las Instituciones de Educación Superior (IES) es más alta en los primeros cuatro semestres (deserción temprana) [2]. Una de las causas académicas de la deserción es la poca orientación profesional y vocacional antes del ingreso a la universidad [3], [4].

En el Sistema para la Prevención de la Deserción en la Educación Superior (Spadies) del Ministerio de Educación de Colombia, el cual consolida y ordena la información que permite hacer seguimiento a las condiciones académicas y socioeconómicas de los estudiantes que han ingresado a la educación superior en el país, se observa que los cuatro primeros semestres son críticos por la gran cantidad de estudiantes que desertan, concretamente el 75 % del total de deserciones, y se identifica que se debe principalmente a sus bajas competencias académicas al momento de ingresar a la universidad, aspectos económicos y aspectos relacionados con orientación socio-ocupacional.

El Ministerio de Educación Nacional de Colombia [3] plantea que la deserción en programas de Ingeniería en Colombia en el primer semestre es del 22,8 %, y la deserción acumulada hasta el quinto semestre es del 46,7 %. El factor académico de mayor riesgo en las Instituciones de Educación Superior se presenta en los programas que pertenecen al área de matemáticas y ciencias naturales [6].

En algunos casos la deserción se debe al desconocimiento que tienen los estudiantes con respecto al programa al que ingresan, al encontrar que no es lo que esperaban o no es de su total agrado [5]. Al analizar los factores a la hora de elegir una carrera, se encuentran [7] los factores internos que forman parte de la personalidad del ser humano, entre ellos habilidades, intereses, actitudes y aptitudes; y los factores externos como la familia y el medio donde se desarrolla.

Como se puede observar, un factor relevante de deserción es la afinidad con la carrera elegida, que se toma como:

... grado de afinidad que existe entre las demandas de las carreras que se ha pensado estudiar y las características personales; es decir una valoración de gustos y habilidades para manejar ciertas cosas, como por ejemplo la lengua hablada, escrita o las matemáticas. Los casos de afinidad entre las características del

estudiante y la carrera que eligió, demuestran que se alcanza un alto rendimiento y un gran disfrute al cursar o ejercer la carrera. [8].

La afinidad académica tiene en cuenta solo la parte de capacidades y aptitudes, y la afinidad personal involucra aspectos más del carácter, las actitudes y gustos del individuo. Al analizar estos aspectos en relación con las cifras de deserción, se encuentra que hay una serie de datos y factores que son desconocidos por un aspirante al momento de seleccionar un programa de educación superior.

Con base en lo anterior, es necesario explorar tecnologías que permitan minimizar el problema de la deserción, tal como la analítica de datos, la cual permite analizar datos y obtener información para mejorar la toma de decisiones. Una de las estrategias de análisis de datos es la minería de datos, que implica la deducción de algoritmos que exploran los datos y el desarrollo de un modelo para descubrir patrones previamente desconocidos [9]. En este aspecto, se han realizado múltiples estudios en universidades, en diferentes áreas del conocimiento, que caracterizan las causas de deserción de los estudiantes, en los que se ha utilizado principalmente minería de datos, como en el proyecto realizado en México [5], donde se identifican causas de deserción escolar en la carrera de Ingeniería en Computación. Estos estudios se han concentrado en la identificación de causas, sin intentar anticiparse a esta situación.

En este proyecto, tomando de base la caracterización de los elementos que influyen en la afinidad del aspirante, se asociaron las características académicas y personales con las características de los programas de Ingeniería del Politécnico Colombiano Jaime Isaza Cadavid (PCJIC), planteando la definición de un modelo de afinidad académica y personal, que incorpora técnicas de análisis de datos. A partir del diseño, se desarrolló un prototipo que permite la validación del modelo de afinidad en el programa de Ingeniería Informática del PCJIC, y la identificación de formas de mejora para trabajos futuros.

Este documento se estructura de la siguiente manera: En la sección 2, se presentan la revisión de los antecedentes; en la sección 3 los materiales y métodos, en la sección 4 los resultados, en la sección 5 las conclusiones y finalmente las referencias.

II. REVISIÓN DE ANTECEDENTES

La Universidad de los Andes (Colombia), en conjunto con el Ministerio de Educación Nacional [10], realizaron la investigación referente a deserción académica que ha tenido más relevancia, cuyo objetivo fue analizar información de las Instituciones de Educación Superior (IES) y determinar las características que son influyentes en cuanto a deserción se refiere, calculando probabilidades de riesgos de estudiantes particulares y de grupos de individuos con características semejantes. La investigación dio lugar al Sistema de Prevención y Análisis de la Deserción en las Instituciones de Educación Superior (Spadies), el cual se alimenta de la información que genera el Instituto Colombiano de Crédito Educativo y Estudios Técnicos en el Exterior (Icetex), el Instituto Colombiano para la Evaluación de la Educación (Icfes) y el Sistema Nacional de Información de la Educación Superior (SNIES). Spadies ha sido de gran utilidad para posteriores investigaciones por la información estadística que brinda acerca de las IES en Colombia.

Algunos estudios [11], [12] emplearon redes neuronales, con el objetivo de determinar causas y patrones de deserción en cada una de sus universidades y casos particulares, y han encontrado en la minería de datos una buena técnica para el estudio de escenarios complejos como la deserción.

Otros estudios [13]-[16] utilizan técnicas de minería de datos con el objetivo de observar e identificar factores que influyen en la deserción académica y en la identificación de patrones de bajo rendimiento. En ellos se encontraron diversos factores que influyen en la deserción, y se recomienda emplear Weka y TaryKDD para realizar estudios con minería de datos.

En un estudio [15] con minería de datos se empleó la metodología CRISP-DM. Su objetivo fue identificar patrones en estudiantes con sus perfiles socioeconómicos y académicos en programas de pregrado de IES, buscando lograr una mayor retención estudiantil. Con la información y los patrones obtenidos se recomendó a las directivas universitarias de ambas IES formular políticas y estrategias relacionadas con los programas de retención estudiantil.

En otro estudio [17] realizado utilizando técnicas de minería de datos e inteligencia artificial, se extrajo información mediante algoritmos de aprendizaje automático. Su objetivo fue diseñar modelos predictivos con la finalidad de integrar al estudiante en su plan de retención en caso de que las características lo indiquen. El estudio muestra lo útiles que son los algoritmos de aprendizaje automático como los árboles de decisión, que ayudan a crear modelos predictivos eficaces a partir de los datos de retención estudiantil acumulados de años anteriores.

En un estudio [18] que emplea técnicas de minería de datos y la metodología CRISP-DM, el objetivo fue analizar las calificaciones de los alumnos de ingeniería de sistemas en las materias Algoritmos y Estructura de Datos para caracterizar los perfiles de los alumnos exitosos y de aquellos que no, debido al bajo rendimiento académico, con el fin de revertir los malos resultados y determinar las variables asociadas. Se estableció que el modelo es una herramienta válida, veraz, de gran utilidad para la gestión académica y muestra que se puede implementar en diversas instituciones para indicar, según el perfil del alumno, la alta probabilidad de fracaso académico, para, con base en ello, tomar medidas que tiendan a reducir el fracaso académico.

Se llevaron a cabo varios estudios [19]-[22] de tipo predictivo para encontrar patrones que inducen a los estudiantes a abandonar sus carreras y para prever cómo será su rendimiento académico, en los que se emplearon diversas técnicas de minería de datos, como árboles de decisión, reglas de asociación y regresión logística.

Otros estudios [18], [23], [24] que utilizaron minería de datos para evaluar el desempeño de alumnos de primer semestre en las materias escogidas, también analizaron la posible deserción para brindarles mayor atención en su proceso. Se determinaron perfiles y la importancia de programas de tutoría en su acompañamiento.

En [25] se presenta un estudio en el que se emplea minería de datos, con la herramienta Weka. Su objetivo fue encontrar un clasificador del rendimiento académico y detectar los patrones que influyen en la deserción de los estudiantes del departamento de Ingeniería e Investigaciones Tecnológicas.

El objetivo del estudio en [26], utilizando minería de datos y reglas de asociación, fue determinar patrones en la evaluación parcial de la materia Álgebra Lineal para estudiantes de primer año, con el fin de conocer el grado de incidencia de los ejercicios en el hecho de que los alumnos aprueben o desapruében el parcial. Se determinó la importancia que tiene la aplicación de la técnica de reglas de asociación, la cual permitió realizar un análisis novedoso de los resultados para el grupo de ejercicios planteados, y se obtuvieron distintas reglas de asociación cuya relevancia y significancia fue cuantificada por indicadores de soporte, confianza y *lift*.

Se presenta una investigación [27] para identificar las causas por las cuales los estudiantes de primer año abandonaban sus estudios. La información se obtuvo mediante cuestionarios y dio como resultado principalmente dos elementos: el primero relacionado con la carrera elegida al matricularse y el segundo relacionado con la frustración de expectativas. Además, se evidenció cómo la conexión de los estudios con sus características y sus habilidades personales y la conexión de los estudios con sus intereses y gustos, disminuyen la probabilidad de que los alumnos de primer año abandonen la carrera. En la investigación se resalta cómo la labor orientadora a la hora de elegir estudios universitarios es primordial para que los estudiantes no presenten una deserción temprana.

En otros estudios, por medio del uso del descubrimiento de conocimiento basado en reglas y árboles de decisión, se han identificado reglas de comportamiento a nivel personal y académico que tienen incidencia en el retiro, como quién financia los estudios y la cantidad de tiempo transcurrido desde que finalizó el colegio e ingresó a la universidad [28], [29].

En los diferentes antecedentes se identificó que los proyectos que estudian la deserción, lo hacen principalmente para caracterizar y encontrar patrones con el fin de prestarles más atención o brindarles tutorías; en la mayoría se tiene en cuenta solo el aspecto académico, no se identificaron trabajos donde se utilicen estos patrones con el fin de orientar al aspirante y evitar una posible deserción.

III. MATERIALES Y MÉTODOS

En esta sección se describe la metodología del proceso para la caracterización, el análisis de datos y el modelo propuesto para determinar la afinidad de los aspirantes con el programa de Ingeniería, particularmente el prototipo realizado con Ingeniería Informática del Politécnico Colombiano Jaime Isaza Cadavid (PCJIC).

A. Caracterización del Aspirante y de los Perfiles de Programas de Ingeniería del PCJIC

Para la caracterización se revisó la literatura, con el fin de identificar en ésta características comunes de los ingenieros [30], [31]. Se analizó la literatura en lo referente a factores de afinidad con la carrera y de orientación vocacional profesional [7] con el propósito de identificar aspectos relevantes, tales como factores internos y externos que inciden en la elección. Asimismo, se analizaron los perfiles de los programas de Ingeniería del PCJIC y se identificaron características comunes y propias para cada programa.

Con el fin de validar la caracterización identificada, se realizó una encuesta, con un nivel

de confianza del 96 %, y se obtuvo una muestra de 104 estudiantes, del total de la población de 4.888 matriculados en ingeniería para el semestre 2018-2, según el Sistema de Indicadores y Estadísticas (SIE) del PCJIC.

B. Análisis de Datos

Para el análisis de datos se utilizó la metodología CRISP-DM [32] que define seis fases: entendimiento del negocio, entendimiento de los datos, preparación de los datos, modelado, evaluación e implementación.

En el entendimiento del negocio y en el de los datos, se revisó la información disponible en el Politécnico que pudiera relacionarse con la caracterización identificada, para conocer las características de los estudiantes actuales y graduados de los programas de Ingeniería.

En el PCJIC se cuenta con datos de la ficha psicosocial que se obtienen al momento de ingresar el estudiante a los seis programas de Ingeniería, se contó con información correspondiente a los periodos 2010-1 a 2017-2; además, con información de tipo personal, familiar, académica, laboral, cívica y de informática. También se tuvo acceso a informes de deserción e información académica de los estudiantes de Ingeniería Informática y de los graduados.

En la fase de preparación de datos, se utilizaron las siguientes herramientas de software: para el proceso de ETL se utilizó la herramienta Pentaho Data Integration, y para la recopilación, integración y almacenamiento de los datos del modelo el motor de base de datos fue PostgreSQL.

Para la fase de modelado se seleccionó la técnica, tomando como base los proyectos de minería de datos consultados en la revisión de literatura; se tuvieron en cuenta las características más apropiadas de acuerdo con el modelo para hacer la comparación entre ellas. En la Tabla 1 se muestran las características de las técnicas evaluadas [33].

Tabla 1. Técnicas de analítica de datos

Técnica / Característica	Clustering	Reglas de asociación	Árboles de decisión
Descriptiva			x
Predictiva	x	x	
Relación entre registros	x		
Agrupamiento de registros		x	x
Clasificación de registros		x	

Fuente: Elaboración propia.

Se eligió una técnica de tipo descriptiva para comprender mejor los datos que se encuentran en las diferentes fuentes, se consideró agrupar y clasificar los estudiantes según sus características, y además se buscó identificar relaciones que no se podrían derivar a simple vista; por lo cual la técnica elegida fue Clustering.

Para la fase de implementación de la técnica de analítica de datos se eligió y utilizó Weka, por ser una plataforma de software libre para el aprendizaje automático y por haber mostrado gran desempeño en los proyectos de analítica de datos consultados en los antecedentes.

C. Modelo de Afinidad

Con base en la caracterización de los aspirantes y de los programas de Ingeniería, y el análisis de la información almacenada en la Institución de los estudiantes y graduados con las técnicas pertinentes de analítica de datos, se procedió a elaborar un algoritmo basado en los patrones encontrados, por medio del cual se obtiene el porcentaje de afinidad con los factores internos y externos y con el programa de Ingeniería Informática.

Este proceso se lleva a cabo mediante un formulario web diseñado para capturar la información del aspirante, y proceder a mostrarle su afinidad con el

programa de Ingeniería, además por ser el modelo validado en ingeniería informática, se incluyeron preguntas sobre características específicas de este programa.

Para obtener los datos de los aspirantes a programas de Ingeniería, a los cuales se les va a presentar la afinidad, se empleó el formulario web, donde ellos ingresan información relacionada con los factores internos y externos que los caracterizan; además, se incluyen las características halladas que debe tener un ingeniero informático, tales como la lógica y el manejo de dispositivos tecnológicos.

IV. RESULTADOS

A. Caracterización del Aspirante y de los Perfiles de Programas de Ingeniería del PCJIC

La caracterización se divide en dos elementos: 1) el aspirante con sus características personales y académicas; 2) los programas de Ingeniería llevados al contexto del Politécnico Colombiano Jaime Isaza Cadavid (PCJIC).

Teniendo en cuenta lo planteado en la literatura [7] y las características definidas en la revisión de antecedentes [30], [31] que debe tener un estudiante de Ingeniería, se identificaron las siguientes: conocimientos sólidos; habilidades y capacidades, como el trabajo en equipo, el autoaprendizaje, el liderazgo, la creatividad, la innovación, la investigación, la comunicación, el dominio de la informática y la adaptación; características de entendimiento y comprensión, como la toma de decisiones, el pensamiento crítico, el análisis y el mejoramiento continuo; y cualidades personales, actitudes y valores, como la ética, la perseverancia, el respeto, el compromiso y la responsabilidad.

En la Tabla 2 se presentan de manera detallada las características personales y académicas identificadas para los profesionales en Ingeniería.

Tabla 2. Características personales y académicas de los estudiantes de programas de Ingeniería

Ingeniero	
Características personales	Desempeñarse de manera efectiva en equipos de trabajo Asumir roles y responsabilidades Comunicarse con efectividad Actuar con espíritu emprendedor Actuar éticamente
Características académicas	Identificar problemas Formular problemas Resolver problemas Aprender de forma continua y autónoma Tener buena formación básica

Fuente: Elaboración propia.

Al revisar los perfiles de los seis programas de Ingeniería del Politécnico Colombiano Jaime Isaza Cadavid (PCJIC), se encontraron cinco características en común para todos, a saber:

- Capacidad de análisis
- Actitud crítica e innovadora
- Capacidad de trabajar en equipo
- Sólida formación básica
- Espíritu investigativo

Al detallar cada programa, se observan características que son particulares, como se ve en la Tabla 3.

Tabla 3. Perfiles de los programas de Ingeniería del Politécnico Colombiano Jaime Isaza Cadavid (PCJIC)

Programa	Características
Ingeniería Informática	Lógica
Ingeniería Civil	Replantear problemas y actualizarse de forma permanente
Ingeniería Agropecuaria	Toma de decisiones y conciencia ambiental

Programa	Características
Ingeniería de Instrumentación y Control	Lógica y liderazgo
Ingeniería de Higiene y Salud Ocupacional	Aprender a aprender y desaprender, comunicación, pensamiento sistémico
Ingeniería de Productividad y Calidad	Liderazgo y toma de decisiones

Fuente: Elaboración propia.

Se realizó la encuesta a los 104 estudiantes con el fin de validar la información obtenida en la caracterización con base en la literatura y en la revisión de los perfiles. En la Figura 1 se ve la distribución de los estudiantes por programa, y en la Figura 2, por nivel académico.

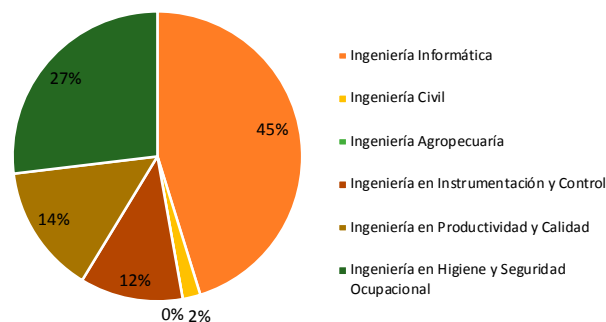


Fig. 1. Clasificación por programas

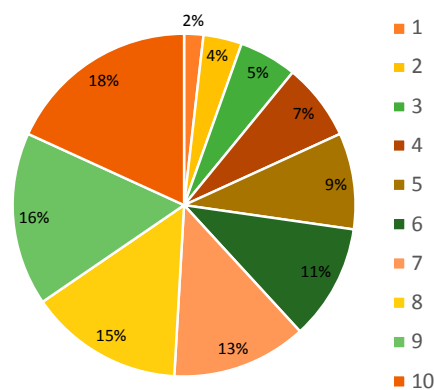


Fig. 2. Clasificación por nivel académico

Se identificó con la encuesta que el 73 % de los estudiantes dicen tener habilidades para trabajar en equipo, el 63 %, para analizar y el 39 %, para investigar; características que fueron encontradas en los perfiles de los programas de Ingeniería del Politécnico Colombiano Jaime Isaza Cadavid (PCJIC).

Además, el 82,7 % dicen ser responsables, el 69,2 % tienen habilidades para resolver problemas, el 39,4 %, para comunicar asertivamente y el 38,5 %, para emprender; características expuestas que la Asociación Iberoamericana de Instituciones de la Enseñanza de la Ingeniería considera necesarias.

Se identificó también que el 62,5 % de los estudiantes que respondieron la encuesta se consideran curiosos, el 61,5 %, perseverantes, el 58,7 % son dedicados, el 51,9 % tienen habilidades para tomar decisiones, el 51 % son abiertos al cambio y el 44,2 % son creativos; las características se obtuvieron de la caracterización del aspirante.

Estos resultados muestran que la caracterización identificada en su mayoría coincide con porcentajes altos en las respuestas de los estudiantes de Ingeniería del PCJIC. Teniendo en cuenta las caracterizaciones anteriores, tanto de los aspirantes como de los programas de Ingeniería, que abarcan las comunes para todos y las particulares, se presenta la asociación en la Figura. 3.



Fig. 3. Asociación de características personales y académicas con los programas de Ingeniería

A partir de la caracterización de los elementos que influyen en la afinidad del aspirante y basado en expertos el modelo se define en la sección posterior.

B. Análisis de Datos

Se ejecutó el método SimpleK-means, gracias a su efectividad [30] y al reconocimiento obtenido en la revisión de antecedentes y la opción Use Training Set, con diferente cantidad de clústeres, y se encontró el mejor desempeño con cuatro grupos, donde la cohesión (distancia de los datos en cada clúster respecto a su centro) fue menor. Se utilizaron 968 instancias y se clasificaron de acuerdo al atributo graduado. Posteriormente, se hizo el perfilamiento de los clústeres, explicando el centroide encontrado por el método SimpleK-means. En la Figura 4 se presenta la conformación de los grupos.

A continuación, la descripción de cada clúster.

Clúster 0. No graduados. Hombres, viven en Medellín, no laboran, los estudios los pagaban sus padres, viven con ambos padres, se transportan en metro, solteros, estrato 2,49, ambos padres trabajan y terminaron secundaria.

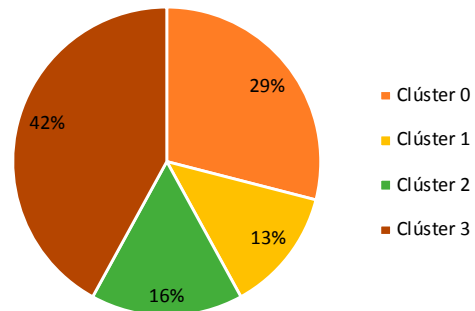


Fig. 4. Conformación de los clústeres

Clúster 1. Graduados. Provenientes de Medellín, estrato 1,96 en promedio, no trabajaban al momento de ingresar a la Universidad, sin hijos ni personas a cargo, sus estudios los pagaban sus padres y no respondieron información de estudio y laboral de sus padres.

Clúster 2. Graduados. Solteros, al momento de ingresar a la Universidad no trabajaban, el gasto de sus estudios era asumido por sus padres, sin hijos ni personas a cargo, su padre trabaja en la industria, viven con parientes, la madre terminó primaria y el papá secundaria, estrato 2,5 en promedio.

Clúster 3. No graduados. Hombres, no provienen de Medellín, solteros, se transportan en metro, viven con la madre, estrato 2,37 en promedio, ambos padres trabajan y terminaron primaria; trabajan y costean sus estudios con recursos propios y perciben entre 1 y 2 salarios mínimos.

Estos resultados indican que los estudiantes graduados se encuentran en el grupo con las siguientes características: al momento de ingresar a la Universidad no trabajaban, el gasto de sus estudios era asumido por sus padres, no tenían hijos ni personas a cargo.

Con base en los datos obtenidos, se procede a construir el modelo de afinidad, se involucran los patrones identificados en un algoritmo que permite calcular el porcentaje de afinidad.

C. Modelo de Afinidad

Con base en la caracterización identificada en los perfiles de los programas de Ingeniería del Politécnico Colombiano Jaime Isaza Cadavid y la encuesta de validación realizada, se construyó el modelo de afinidad académica y personal, el cual tiene como entradas las fuentes de datos mencionadas anteriormente, a las cuales se les realizaron procesos de ETL, luego se almacenaron en una base de datos y posteriormente se agruparon las salidas, tal como se muestra en la Figura 5.

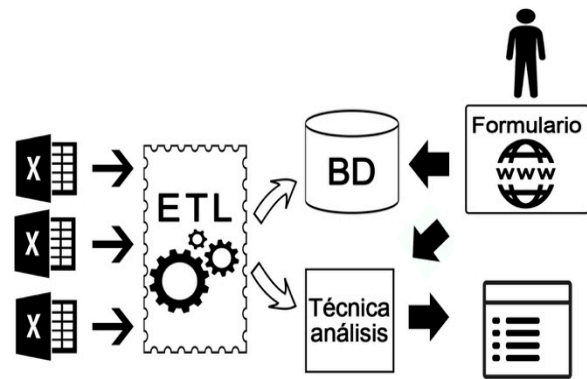


Fig. 5. Modelo de afinidad académica y personal

El modelo se implementó para ser validado en el programa de Ingeniería Informática. En la Figura 6, se muestra la arquitectura del prototipo desarrollado y sus componentes.

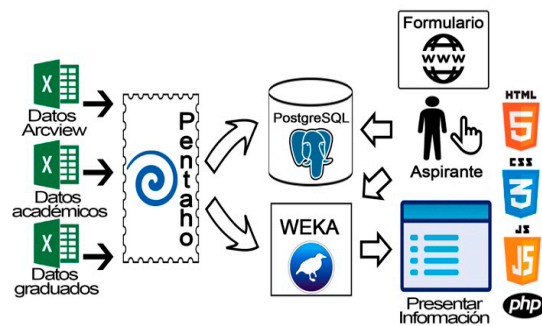


Fig. 6. Arquitectura del prototipo

Se diseñó un formulario para recolectar la información solicitada al aspirante, los datos incluyen información que se identificó en la caracterización y factores internos y externos que influyen en su decisión; se incluyen, además, las características halladas que debe tener un ingeniero informático (ver Figura 7).

Fig. 7. Software web – recolección de datos

En la Figura 8 se puede ver un ejemplo de las preguntas de lógica incluidas y en la Figura 9, de las preguntas sobre conocimientos de informática.

Fig. 8. Formulario Web – preguntas de lógica

Fig. 9. Software web – conocimientos de informática

Con base en la información suministrada por el aspirante, la aplicación realiza el cálculo de la afinidad con los factores externos e internos y con los aspectos propios de la Ingeniería Informática.

Fig. 10. Software web – presentación de resultados

La comprobación se hizo tomando algunos graduados, diez específicamente, a quienes, con base en la información suministrada al ingreso, se les preguntó de nuevo información socioeconómica, de lógica y de conocimientos de informática, y se obtuvo para cada uno el porcentaje de afinidad con el programa, que mostró un alto grado.

V. CONCLUSIONES

Al realizar el proceso de ETL se encontró que la calidad de la información era baja, debido a que el formulario de ingreso utilizado en el PCJIC contiene demasiadas preguntas abiertas y no obligatorias, fue necesario aplicar varios procesos de transformación para mejorar la calidad de los datos. Para aumentar la calidad de la información, sería necesario hacer seguimiento a la información solicitada por la Oficina de Admisiones.

La afinidad con el programa que se desea estudiar depende de múltiples factores, no solo se debe tener en cuenta el aspecto académico, sino prestar atención a los gustos y motivaciones en el momento de elegir una carrera para evitar decepciones y posibles retiros.

La técnica de análisis de datos seleccionada permitió identificar patrones en los datos para así establecer el nivel de afinidad con los graduados del programa, incluyendo factores personales que no se identifican fácilmente a simple vista. Sin embargo, como trabajo futuro sería útil probar otras técnicas de análisis de datos para contrastar resultados, incluso que se pueda aprender de los datos ingresados.

Como trabajo futuro se propone también validar el prototipo en otros programas de Ingeniería del Politécnico Colombiano Jaime Isaza Cadavid (PCJIC), para incorporar las características específicas de cada uno de ellos.

REFERENCIAS

- [1] D. M. M. Bernal, “La deserción escolar: Un problema de carácter social”, *In Vestigium Ire*, N.º 6, 2013.
- [2] Ministerio de Educación Nacional de Colombia, *Guía para la implementación del Modelo de Gestión de Permanencia y Graduación Estudiantil en Instituciones de Educación Superior*. Bogotá: Imprenta Nacional, 2015. Disponible en: http://www.mineducacion.gov.co/1759/articles-356272_recurso.pdf
- [3] Observatorio de Educación Superior de Medellín (ODES), “Deserción en la Educación Superior, Boletín N.º 5, Medellín: Sapiencia, julio 2017. [En línea]. Disponible en: http://www.sapiencia.gov.co/wp-content/uploads/2017/07/BOLETIN_ODES_DESERCION_EN_LA_EDUCACION_SUPERIOR.pdf
- [4] Ministerio de Educación Nacional de Colombia (Mineducación), *Deserción estudiantil en la educación superior colombiana*. Bogotá: Mineducación, 2010.
- [5] J. A. Mendiola, R. V. Rosas, J. A. A. Velásquez, R. A. Eleuterio y J. M. Romero, “Análisis de deserción escolar con minería de datos”, *Research in Computing Science*, vol. 93, pp. 71-82, 2015.
- [6] Universidad de los Andes, Facultad de Economía, Informe mensual sobre el soporte técnico y avance del contrato para garantizar la alimentación, consolidación, validación y uso de la información del SPADIES. Bogotá: Universidad de los Andes, febrero 2014.
- [7] Z. B. Salinas, “La orientación vocacional y profesional en la selección de carreras”, *Ciencia Unemi*, vol. 4, N.º 6, pp. 97-101, 2015.
- [8] V. V. Salmerón, *Orientación profesional: un enfoque sistémico*. 2.ª ed., vol. 1. México: Prentice Hall, 2004.
- [9] O. Maimon and L. Rokach, *Data Mining and Knowledge Discovery Handbook*. Springer, Heidelberg, 2005.
- [10] I. Montes, P. Almonacid, S. Gómez, F. Zuluaga y E. Tamayo, *Análisis de la deserción estudiantil en los programas de pregrado de la Universidad Eafit*. Serie Cuadernos de Investigación. Medellín: Ed. Universidad Eafit, abril 2010.
- [11] J. E. Gutiérrez, “Descubrimiento de conocimientos en la base de datos académica de la Universidad Autónoma de Manizales aplicando redes neuronales”, Universidad Autónoma de Manizales, Colombia, 2012.

- [12] D. del Alcázar, E. Silva y X. Fernández, “Sistema inteligente para perfilar la deserción en estudiantes universitarios de carreras técnicas”, en *Deserción, calidad y reforma universitaria. Apuntes para el debate*, F. Cevallos. Ed. Cuadernos del contrato social por la educación - Ecuador, N.º 10, pp. 50-65, Quito, junio 2014.
- [13] K. Azoumana, “Análisis de la deserción estudiantil en la Universidad Simón Bolívar, Facultad Ingeniería de Sistemas, con técnicas de minería de datos”, *Pensamiento Americano*, vol. 6, N.º 10, 2013.
- [14] W. Fontalvo, M. P. Castillo y S. Polo, “Análisis comparativo entre las características más relevantes de deserción estudiantil en el programa de Ingeniería Industrial de la Universidad Autónoma del Caribe. Estudiantes activos en el periodo 2013-01 y desertores académicos de los periodos 2011-01 a 2012-02”, *Escenarios*, vol. 12, N.º 1, pp. 96-104, 2014.
- [15] R. Timarán, J. Jiménez, R. T. Pereira y J. J. Toledo, “Detección de patrones de deserción estudiantil en programas de pregrado de instituciones de educación superior con CRISP-DM”, en *Congreso Iberoamericano de Ciencia, Tecnología, Innovación y Educación*, Buenos Aires, 2014, pp. 1-19.
- [16] K. B. Eckert y R. Suénaga, “Análisis de deserción-permanencia de estudiantes universitarios utilizando técnica de clasificación en minería de datos”, *Formación Universitaria*, vol. 8, N.º 5, pp. 03-12, 2015.
- [17] S. H. Lin, “Data mining for student retention management”, *J. Comput. Sci. in Colleges*, vol. 27, N.º 4, pp. 92-99, 2012.
- [18] D. L. La Red, M. Karanik, M. Giovannini y N. Pinto, “Perfiles de rendimiento académico: un modelo basado en minería de datos”, *Campus Virtuales*, vol. 4, N.º 1, pp. 12-30, 2015.
- [19] Y. K. Amaya, E. Barrientos y D. J. Heredia, “Modelo predictivo de deserción estudiantil utilizando técnicas de minería de datos”, 2010 [En línea]. Disponible en: <https://documentos.redclara.net/bitstream/10786/759/1/124-22-3-2014-Modelo%20predictivo%20de%20deserci%C3%B3n%20estudiantil%20utilizando%20t%C3%A9cnicas%20de%20miner%C3%ADa%20de%20datos.pdf>.
- [20] H. C. Ahumada, H. Dip, C. G. Herrera y J. C. Leguizamón, “Minería de datos para un sistema de alerta temprana de deserción en carreras de ingeniería”, en *XVII Workshop de Investigadores en Ciencias de la Computación*, Salta, 2015.
- [21] S. Celis, L. P. Poblete, J. Villanueva y R. Weber, “Un modelo analítico para la predicción del rendimiento académico de estudiantes de ingeniería”, *Rev. Ing. Sistemas*, vol. XXIX, 2015.
- [22] E. A. Porcel, G. N. Dapozo y M. V. López, “Predicción del rendimiento académico de alumnos de primer año de la FACENA (UNNE) en función de su caracterización socioeducativa”, *Rev. Electrónica de Investigación Educativa*, vol. 12, N.º 2, pp. 1-21, 2010.
- [23] E. I. García y P. M. Mora, “Creación de un modelo de predicción del desempeño académico de los alumnos de la Facultad de Ingeniería de La UNAM en el primer semestre”, Universidad Nacional Autónoma de México (UNAM), 2011.

- [24] M. Gil, M. Soria, J. Mosqueda, R. Sánchez, I. Pérez y N. Reyes, “Detección de la probabilidad de deserción en alumnos de nuevo ingreso en base a factores de riesgo utilizando redes neuronales artificiales”, *Tecnologías y Aprendizaje. Avances en Iberoamérica*, vol. 2, p. 52.
- [25] O. M. Sposito, M. Etcheverry, M. Ryckeboer y J. Bossero, “Aplicación de técnicas de minería de datos para la evaluación del rendimiento académico y la deserción estudiantil”, en *Novena Conferencia Iberoamericana en Sistemas, Cibernética e Informática*, CISCI, vol. 29, 2010.
- [26] H. C. Ahumada y C. Herrera, “Utilización de minería de datos en el análisis de ejercicios de una evaluación de álgebra lineal”, *ReTyCA (Revista de Tecnología y Ciencias Aplicadas)*, vol. 1, pp. 1-6, 2016.
- [27] M. Álvarez, P. Figuera y M. Torrado, “La problemática de la transición Bachillerato-Universidad en la Universidad de Barcelona”, *REOP-Revista Española de Orientación y Psicopedagogía*, vol. 22, N.º 1, pp. 15-27, 2011.
- [28] H. Kuna, R. García and F. Villatoro, “Pattern discovery in university students’ desertion based on data mining”, in *Proc. IV Meeting on Dynamics of Social and Economic Systems*, 2010, pp. 275-285.
- [29] D. Vila, S. Cisneros, P. Granda, C. Ortega, M. Posso-Yépez, I. García-Santillán, “Detection of desertion patterns in university students using data mining techniques: A case study”, in *Technology Trends. CITT 2018. Communications in Computer and Information Science*, M. Botto-Tobar, G. Pizarro, M. Zúñiga-Prieto, M. D’Armas, M. Zúñiga-Sánchez, eds., vol. 895. Germany: Springer, Cham, 2018, pp. 420-429.
- [30] E. Blanco Rivero, “Perfil del ingeniero colombiano para el 2020. Education, Innovation, Technology and Practice”, en *5th Latin American and Caribbean Conference for Engineering and Technology*, Tampico, México, 2007.
- [31] R. G. Lerena, *Competencias y perfil del Ingeniero Iberoamericano, formación de profesores y desarrollo tecnológico e innovación*. Bogotá, Colombia: ASIBEI, 2016.
- [32] S. Sharma, K. M. Osei-Bryson, G. M. Kasper, “Evaluation of an integrated Knowledge Discovery and Data Mining process model”, *Expert Systems with Applications*, vol. 39, pp. 11335-11348, 2012. <http://doi.org/10.1016/j.eswa.2012.02.044>
- [33] J. Han, J. Pei, J. and M. Kamber, *Data Mining: Concepts and Techniques*. Elsevier, 2011.

Capítulo IX

Factores que Inciden en el Funcionamiento de una PMO en las Empresas de TI: Una Revisión de Literatura

Adriana Xiomara Reyes Gamboa - axreyes@elpoli.edu.co

Sara Cristina Vera Sepúlveda - sara_vera54182@elpoli.edu.co

Politécnico Colombiano Jaime Isaza Cadavid, Medellín, Colombia

Darío Enrique Soto Durán - dsoto@tdea.edu.co

Docente-investigador, Tecnológico de Antioquia - I. U., Medellín, Colombia

I. INTRODUCCIÓN

Una Oficina de Gestión de Proyectos (Project Management Office —PMO—) se define como una estructura organizativa que estandariza los procesos de gobierno relacionados con el proyecto y facilita el intercambio de recursos, metodologías, herramientas y técnicas [1]. En [2] se define como una estructura organizativa creada para promover y mejorar la práctica de la gestión de proyectos, mediante la adopción de metodologías apropiadas para lograr altos niveles de eficiencia y eficacia. Las organizaciones de Tecnología de la Información (TI) deben llevar a cabo cambios significativos e implementar nuevas prácticas más ajustadas a la realidad, incluido el uso de enfoques de gestión de proyectos y beneficios, buscando un mejor uso y control de los recursos y capacidades existentes [3].

Hay diferentes versiones de factores que influyen en el funcionamiento de una PMO, por ejemplo, para [4], la combinación de una excelente investigación y la gestión profesional de proyectos es un factor de éxito cuando se maneja una gran cartera de proyectos complejos. Para [5], el éxito radica en la comunicación que se tiene con la alta dirección de la empresa de manera continua y exitosa y es lo que asegura la validez de la estrategia. También se habla acerca de la gestión de proyectos como esencial en el contexto del desarrollo de proyectos exitosos, es transversal y tiene aplicaciones en muchos sectores [6]. Los mismos autores afirman que es particularmente cierto en proyectos grandes, donde

la necesidad de una estructura de administración de proyectos competente se hace más evidente y realmente indudable debido a la complejidad involucrada, y más aún en los proyectos del sector TI que incluyen características particulares como: alta incertidumbre, alta complejidad y ambientes de permanente cambio, por lo anterior es que se hace necesario identificar los factores que inciden en el éxito de las PMO para aplicarse al contexto de las empresas de TI.

En general, en las organizaciones de tecnologías de información se trabaja con marcos de gobierno de TI, sin embargo, a pesar de que existe relación con la gestión de proyectos se omiten de las funciones de PMO dentro de dichos marcos [7]. Una PMO y la estructura organizacional proyectada son aspectos positivos y significativos en la alineación de los proyectos de TI con los modelos de negocio [8]. Sin embargo, en este aspecto se requeriría un análisis más detallado de los impactos de este tipo de metodologías en las empresas de TI, debido a que lamentablemente, aunque la existencia de una PMO puede tener un impacto en el rendimiento de la organización, el desinterés por parte de los equipos de gestión para establecerla sigue siendo una tendencia frecuente en múltiples industrias [9]. En [9] se indica que se requieren prácticas tales como el desarrollo de métricas de rendimiento para medir el valor de la PMO en IT, la estructuración de éstas basadas en una práctica de gestión de servicios y un modelo conceptual propuesto específico para la industria.

El presente artículo se encuentra organizado en tres secciones: la primera describe el protocolo que orienta la revisión de la literatura, la segunda corresponde a los resultados y su respectivo análisis y, por último, se presentan las conclusiones.

II. MATERIALES Y MÉTODOS

A. Materiales

La literatura que existe sobre el tema provee una gran variedad de artículos que hablan sobre los factores que influyen en el funcionamiento de una PMO.

B. Métodos

La revisión de literatura se realiza bajo el protocolo de revisión Kitchenham (2007). El objetivo de este protocolo es apoyar en esta labor a los investigadores para evitar sesgos en la realización de este proceso. El protocolo plantea tres fases:

- *Planificación:* Se refiere a las actividades previas a la revisión, para definir las preguntas de investigación, los criterios de inclusión y exclusión, las fuentes de información, la cadena de búsqueda, y los procedimientos de categorización.
- *Ejecución:* Se refiere a la búsqueda y la selección de los estudios con el fin de extraer los datos y sintetizarlos.
- *Resultados:* Es la fase final. Tiene por objeto redactar los resultados de acuerdo con las preguntas de investigación planteadas. Los resultados derivados de la revisión presentan un análisis crítico de los autores.

III. RESULTADOS

En la Tabla 1 se presentan las preguntas que se propone resolver con la revisión de literatura, así como la justificación de su planteamiento.

Tabla 1. Preguntas de investigación y su justificación

N.º Pregunta de investigación	Justificación de la pregunta
PI1 ¿Cuándo y dónde fue publicado el estudio?	El tema de investigación es amplio y pertinente en las disciplinas de ingeniería y la gestión de proyectos. Esta pregunta de investigación tiene como propósito conocer y comprender las fuentes de las publicaciones específicas para la temática y cuándo han sido publicadas.
PI2 ¿Cuáles son los factores positivos que influyen en el funcionamiento de una PMO?	Se formula esta pregunta de investigación para establecer las categorías donde se identifican factores positivos en el funcionamiento de una PMO.
PI3 ¿Cuáles son los factores negativos que influyen en el funcionamiento de una PMO?	Se formula esta pregunta de investigación para establecer las categorías donde se identifican factores negativos en el funcionamiento de una PMO.

Fuente: Elaboración propia.

Criterios de inclusión y exclusión: La selección de criterios está organizada en un criterio de inclusión (CI) y cuatro criterios de exclusión (CE). El criterio de inclusión es: (CI1) El estudio analiza los factores que inciden en el funcionamiento de una PMO. Los criterios de exclusión son: (CE1) El estudio es un artículo corto. (CE2) El estudio tiene un tiempo de publicación mayor a 5 años. (CE3) La publicación es una versión anterior de otro estudio publicado. (CE4) La publicación es un artículo producto de talleres y tutoriales.

Fuentes de búsqueda: La revisión se enfocó en la base de datos Science Direct (<http://www.sciencedirect.com>).

Palabras clave y cadena de búsqueda: La búsqueda de los estudios se realizó teniendo en cuenta el área de interés de la revisión y construyendo la cadena de búsqueda con las palabras claves. La cadena de búsqueda definida es (“success” OR “issue”) AND (“PMO” OR “project” OR “project management”).

Validación: Antes de realizar la búsqueda de la información, el protocolo de revisión se validó evaluando la pertinencia de las palabras claves y la cadena de búsqueda, con una pequeña muestra de publicaciones.

Extracción de la información y síntesis: El proceso de búsqueda consideró publicaciones de antes del mes de mayo de 2019. La última búsqueda se realizó en mayo de 2019. Como resultado de la búsqueda se obtuvo un total de 31 publicaciones. Se aplicaron los criterios de selección (de inclusión y de exclusión) sobre fecha, título, resumen y palabras clave, lo que dio como resultado 20 documentos (reducción aproximadamente del 36 %), y quedó el siguiente listado:

Tabla 2. Listado de artículos seleccionados

1	PMO implementation experiences in companies of Medellin city
2	The role of the Project Management Office (PMO) in product lifecycle management: A case study in the defense industry
3	Importance of Managing PMO in Croatian PM Market
4	Improving Project Success: A Case Study Using Benefits and Project Management
5	A Performance Evaluation Model for Project Management Office Based on a Multicriteria Approach
6	Integrated Framework for Project Management Office Evaluation

7	PMO as a key ingredient of public sector projects' success – position paper
8	Project Management Office Models – a review
9	The Project Office as Project Management Support in Complex Environments
10	Projects' issue management
11	Relationships between a project management methodology and project success in different project governance contexts
12	Project management office in non-governmental organizations: an ex post facto study
13	Project complexity and team-level absorptive capacity as drivers of project management performance
14	Why and how do project management offices change? A structural analysis approach
15	Software project management in high maturity: A systematic literature mapping
16	Benefits Realisation Management and its influence on project success and on the execution of business strategies
17	Identifying organizational variables affecting project management office characteristics and analyzing their correlations in the Iranian project-oriented organizations of the construction industry
18	Caracterización de los stakeholders que se relacionan con las Oficinas de Gerencia de Proyectos-Fundamento del ecosistema para la gestión de proyectos sostenibles
19	PMO CONCEPTUALIZATION FOR ENGINEERING AND CONSTRUCTION BUSINESSES
20	EXPLORING THE DYNAMICS OF PROJECT MANAGEMENT OFFICE AND PORTFOLIO MANAGEMENT CO-EVOLUTION: A ROUTINE LENS

Fuente: Elaboración propia.

¿Cuándo y dónde fue publicado el estudio? (PI1). Para dar respuesta a la Pregunta de Investigación 1, las publicaciones se caracterizaron teniendo en cuenta los siguientes criterios:

- Año de publicación (Figura 1).
- País de la publicación (Figura 2).

En la Figura 1 se puede observar que el tema en referencia ha tenido relevancia en los últimos años (100 % de las publicaciones menor a 5 años).

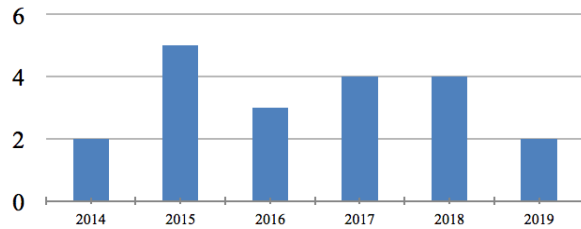


Fig. 1. Cantidad de artículos por año

Fuente: Elaboración propia.

En la Figura 2 se puede concluir que, frente a la temática, los países con mayor número de publicaciones son Brasil, Portugal y UK (45 % de los artículos).

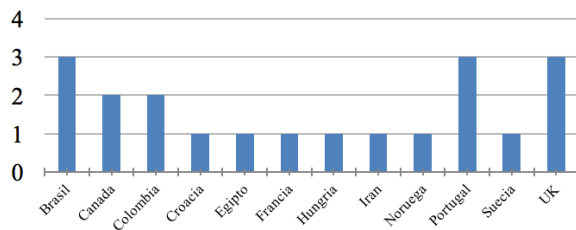


Fig. 2. Cantidad de artículos por país

Fuente: Elaboración propia.

¿Cuáles son los factores positivos que influyen en el funcionamiento de una *PMO*? (PI2). De acuerdo con la revisión, se establecieron los siguientes factores que influyen de manera positiva en el funcionamiento de una *PMO* (Figura 3): El apoyo a las decisiones de la alta dirección, la alta gestión de las comunicaciones y la aplicación metodológica de gestión de proyectos.

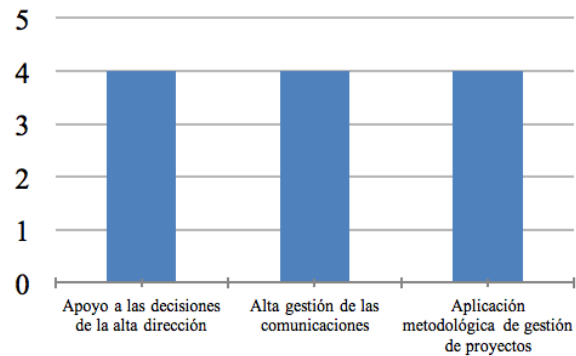


Fig. 3. Factores positivos en el funcionamiento de una *PMO*

Fuente: Elaboración propia.

¿Cuáles son los factores negativos que influyen en el funcionamiento de una *PMO*? (PI3). De acuerdo con la revisión, se establecieron los siguientes factores que influyen de manera negativa en el funcionamiento de una *PMO* (Figura 4): Dificil aplicación de la teoría a la práctica, poco entendimiento de una metodología de gestión de proyectos, y la complejidad e inconvenientes en los proyectos.

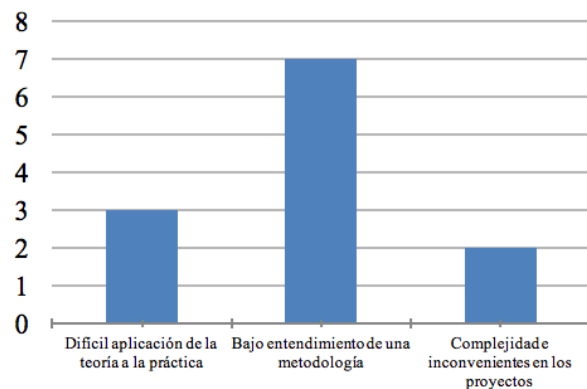


Fig. 4. Factores negativos en el funcionamiento de una *PMO*

Fuente: Elaboración propia.

IV. DISCUSIÓN

Para responder a la pregunta de investigación 2, acerca de los factores positivos que influyen en el funcionamiento de una PMO, se pueden analizar los siguientes:

Apoyo a las decisiones de la alta dirección: En esta categoría se agrupan los artículos 2, 4, 5 y 17, que hacen referencia a que un factor positivo es la medida en la que la PMO ayuda a la alta dirección a tomar decisiones y a obtener beneficios de la gestión de proyectos. Los roles y funciones de una PMO se definen de acuerdo con los objetivos estratégicos de la organización y, en general, deben cumplir con las expectativas de la gerencia superior y los gerentes de proyectos [10]. La gestión de beneficios le permite al patrocinador tener la información correcta para los estudios de viabilidad de inversión y muestra claramente el camino para la entrega de los beneficios esperados de la PMO [11].

Alta gestión de las comunicaciones: En esta categoría se agrupan los artículos 3, 6, 13 y 18, que hacen referencia a que la efectividad en las comunicaciones es otro factor de éxito dentro de la gestión de proyectos. El artículo 3 plantea que la comunicación con la alta gerencia de la compañía es continua y exitosa, lo que solo mejora la importancia del soporte de la alta gerencia y la comunicación general [5].

Aplicación metodológica de gestión de proyectos: En esta categoría se agrupan los artículos 3, 6, 8 y 15, que hacen referencia a que otro factor positivo está relacionado con que la PMO tiene más éxito siempre y cuando aplique una metodología y haga que la organización se apropie de ella. Los autores del artículo 6 afirman que una PMO define y mantiene procesos internos y administra proyectos y portafolios relacionados con procesos institucionales. Estos procesos se esfuerzan por introducir una economía de repetición en la ejecución de proyectos y reducir el riesgo del proyecto a través de la práctica común y el aseguramiento de la calidad [12]. A la vez que

en el artículo 8 se menciona que el crecimiento más reciente parece haber sido impulsado por el deseo de obtener un mejor control de los riesgos del proyecto, estandarizar el uso de metodologías, herramientas y técnicas de gestión de proyectos, mejorar el monitoreo del desempeño del proyecto y administrar y difundir el conocimiento de una gestión de proyectos sólida y práctica [2].

Para responder a la pregunta de investigación 3, acerca de los factores negativos que influyen en el funcionamiento de una PMO, se puede analizar lo siguiente:

Difícil aplicación de la teoría a la práctica: En esta categoría se agrupan los artículos 1, 19 y 20, que hacen referencia a que un factor negativo es que las metodologías de gestión de proyectos no necesariamente sirven para aplicarse textualmente o según la teoría en las empresas, por eso se deben analizar metodologías apropiadas para los proyectos TI. Según los autores del artículo 1, la investigación realizada entre empresas de diversas actividades económicas mostró que no es posible estandarizar una única metodología para todas ellas, sino que es necesario primero entender el ADN de cada organización para hacer la adaptación de una metodología que le corresponda, ya que existe una concepción equivocada acerca de la implementación de las metodologías de proyectos, al creer que pueden seguirse como un recetario [13]. De hecho, los roles y funciones de la PMO, aunque están estandarizados y claramente definidos en la literatura, varían en la práctica [14].

Poco entendimiento de una metodología de gestión de proyectos: En esta categoría se agrupan los artículos 7, 8, 11, 14, 16, 19 y 20, en ellos se refiere que otro factor negativo tiene que ver con que pocas empresas entienden el concepto o la teoría de una metodología de gestión por proyectos. Los autores del artículo 7 indican que, si bien las PMO ahora son una característica destacada de la gestión de proyectos organizativos, la lógica subyacente que conduce a su implementación o renovación aún no se comprende por completo

[6]. Los proyectos aún no alcanzan sus objetivos y todavía falta una cuantificación del impacto de las metodologías en el éxito del proyecto, lo que se evidencia más en los proyectos de TI por sus características particulares [15].

Complejidad e inconvenientes en los proyectos:

En esta categoría se agrupan los artículos 9 y 10, que hacen referencia a que otro factor negativo podría ser la alta complejidad de los proyectos y los inconvenientes que estos acarrearán. Los autores afirman que los tipos y grados de complejidad y la fase del proyecto afectan la gestión del proyecto de diferentes maneras [4].

V. CONCLUSIONES

Uno de los propósitos fue lograr que la mayoría de artículos seleccionados en la revisión bibliográfica fueran publicaciones recientes, lo que permite afirmar que en el artículo se trató una temática actual y vigente. Adicionalmente, se evidenció que los países con más publicaciones al respecto fueron Brasil, Portugal y Canadá, y se encontró también mucha diversidad en los países de origen de las publicaciones, lo que indica que es un tema de interés mundial.

Y, finalmente, para el funcionamiento de una PMO, los factores positivos encontrados fueron: el apoyo a las decisiones de la alta dirección, la alta gestión de las comunicaciones y la aplicación metodológica de gestión de proyectos; los factores negativos encontrados fueron: difícil aplicación de la teoría a la práctica, poco entendimiento de una metodología de gestión de proyectos por parte de las organizaciones y la complejidad y los inconvenientes en los proyectos.

Estos factores identificados permiten ser el punto de partida para definir los que se deben tener en cuenta en la implementación de las PMO en empresas de TI, de manera que se puedan mitigar los factores asociados a la incertidumbre, los frecuentes cambios y la alta complejidad de los proyectos de TI. Un trabajo futuro es incluir investigaciones

que profundicen en impactos de las PMO en las empresas de TI, de acuerdo con necesidades por resolver, tales como métricas, estructuración y conceptos más específicos de esta industria.

REFERENCIAS

- [1] S. Paton & B. Andrew, "Project Management Office (PMO) in product lifecycle management: A case study in the defence industry". *Int. J. Prod. Econ.*, N.º 208, pp. 43-52, September 2016. <https://doi.org/10.1016/J.IJPE.2018.11.002>
- [2] A. Monteiro, V. Santos, & J. Varajão, "Project Management Office models - A review", *Procedia Comp. Sci.*, N.º 100, pp. 1085-1094, 2016. <https://doi.org/10.1016/j.procs.2016.09.254>
- [3] J. Gomes, & M. Romão, "Improving Project Success: A case study using Benefits and Project Management", *Procedia Comp. Sci.*, N.º 100, pp. 489-497, 2016. <https://doi.org/10.1016/j.procs.2016.09.187>
- [4] G. Widforss, & M. Rosqvist, "The Project Office as Project Management Support in Complex Environments", *Procedia Comp. Sci.*, N.º 64, pp. 764-770, 2015. <https://doi.org/10.1016/j.procs.2015.08.626>
- [5] G. Blažević, S. Mišić, & M. Šimac, Importance of Managing PMO in Croatian PM Market. *Procedia - Social and Behavioral Sciences*, N.º 119, pp. 949-956, 2014. <https://doi.org/10.1016/j.sbspro.2014.03.106>
- [6] V. Santos, & J. Varajão, PMO as a Key Ingredient of Public Sector Projects' Success - Position Paper. *Procedia Comp. Sci.*, N.º 64, pp. 1190-1199, 2015. <https://doi.org/10.1016/j.procs.2015.08.546>
- [7] H. Taylor, J. P., Woelfer, & E. "Artman, Information Technology Governance in Practice: A Project Management Office's Use of Early Risk Assessment as a Relational

- Mechanism”, *Int. J. Inf. Technol. Proj. Manage.*, vol. 3, N.º 3, pp. 14-30, 2012. <https://doi.org/10.4018/jitpm.2012070102>
- [8] F. B. Lopez, “Information Technology Project Management Office: Exploring the reasons influencing the adoption of IT PMOs for organizational success”, *ProQuest Dissertations and Theses*, N.º 168, 2016. Disponible en: https://library.capella.edu/login?url=http://search.proquest.com/docview/1809109612?accountid=27965%5Cnhttp://wv91q51d3p.search.serialssolution.com?ctx_ver=Z39.88-2004&ctx_enc=info:ofi/enc:UTF-8&rft_id=info:sid/ProQuest+Dissertations+%26+Theses+Global&rft
- [9] L. D. Siqueira, & S. F. Crispim, “Alinhamento dos projetos de TI aos modelos de negócio das organizações”, *Gestao e Producao*, vol. 21, N.º 3, pp. 621-634, 2014. <https://doi.org/10.1590/0104-530X294>
- [10] T. G. A. Viglioni, J. A. O. G., Cunha, & H. P. Moura, “A Performance Evaluation Model for Project Management Office Based on a Multicriteria Approach”, *Procedia Comp. Sci.*, N.º 100, pp. 955-962, 2016. <https://doi.org/10.1016/j.procs.2016.09.257>
- [11] S. Paton, & B. Andrew, “The role of the Project Management Institute”, en *PMBOK Guide Sixth Edition*, 2017.
- [12] I. Szalay, Á. Kovács, & Z. Sebestyén, “Integrated Framework for Project Management Office Evaluation”, *Procedia Eng.*, N.º 196, pp. 578-584, jun. 2017. <https://doi.org/10.1016/j.proeng.2017.08.033>
- [13] C. M. Betancur, I., Pinzón, y J. S. Posada, “Experiencias de implementación de PMO en empresas de la ciudad de Medellín”, *Rev. EIA. Esc. Ing. Antioq.*, N.º 21, ene.-jun. 2014.
- [14] C. S., Oliveira, A. Tereso, & G. Fernandes, “PMO Conceptualization for Engineering and Construction Businesses”, *Procedia Comp. Sci.*, N.º 121, pp. 592-599, 2017.
- [15] A. Lacruz, & E. Cunha, “Project management office in non-governmental organizations: an *ex post facto* study”, *Revista de Gestão*, vol. 25, N.º 2, pp. 212-227, 2018. <https://doi.org/10.1108/REGE-03-2018-033>

Capítulo X

Estimación de Precipitación y Evapotranspiración Utilizando Sistemas de Información Geográfica a partir de Datos Obtenidos en Estaciones Meteorológicas Ubicadas en la Cuenca del Río Negro

Mauricio Escobar Tobón - mauricio_escobar54181@elpoli.edu.co

Sandra Patricia Mateus - spmateus@elpoli.edu.co

Juan Carlos Valdés - jcvaldes@elpoli.edu.co

Politécnico Colombiano Jaime Isaza Cadavid, Medellín, Colombia

I. INTRODUCCIÓN

Conocer el comportamiento hidrológico de una región, especialmente el asociado a la lluvia, permite empezar a entender fenómenos de crecientes que pueden estar fuera de lo normal, también ayuda al entendimiento de los efectos erosivos de la lluvia y las escorrentías [1].

Para la realización de estos procesos se hace necesario la utilización de herramientas que agilicen el análisis y ayuden a la actualización de la información. Es aquí donde los Sistemas de Información Geográfica (SIG) toman importancia, porque su mayor fortaleza es la gran capacidad de trabajar con bases de datos, lo cual la convierte en una herramienta especializada para el análisis y la mapeación de la información hidrológica. Además, en la actualidad, se cuenta con gran cantidad y calidad de información con respecto a variables como, precipitación, temperatura, evapotranspiración, entre otras, esto, unido con las fortalezas de los SIG, crea un escenario propicio de estudio del comportamiento hidrológico de un sitio de interés [2].

Por eso en este artículo se utilizan las herramientas de los SIG para realizar la delimitación de la cuenca del río Negro, a partir de las curvas de nivel obtenidas de estudios realizados por la Corporación Ambiental de las Cuencas de los ríos

Negro y Nare (Cornare). Se hicieron análisis de precipitación, evapotranspiración y temperatura, tomando los datos de las estaciones climatológicas del Instituto de Hidrología, Meteorología y Estudios Ambientales (Ideam) junto con las herramientas de análisis y manejo de datos de los SIG y se presentan los mapas con la información procesada y analizada.

El artículo se estructura de la siguiente forma: en la sesión 2 se presenta el área de estudio; en la sesión 3 se describe la metodología; en la sesión 4 se plantean los resultados, y, por último, se encuentran las conclusiones.

II. REVISIÓN DE LITERATURA

A. Ciclo hidrológico

La evaluación de los recursos hídricos requiere una correcta estimación del balance hidrológico o de la repartición de la precipitación entre evapotranspiración, escorrentía y recarga de acuíferos.

La ecuación de balance hidrológico es la ley más importante en Hidrología, y aunque su expresión es muy simple, la cuantificación de sus términos es normalmente complicada, debido a la falta de medidas directas y a la variación espacial de la evapotranspiración, de las pérdidas profundas y de los cambios del agua almacenada en una cuenca [3].

Para solucionar estas dificultades, generalmente se admiten dos supuestos: el primero es que la cuenca es impermeable, por tal motivo las pérdidas profundas son despreciables; y el segundo admite que las variaciones del agua almacenada en la cuenca son despreciables para un período suficientemente largo; con estos supuestos se logra realizar un acercamiento al balance hidrológico.

Variables involucradas en el estudio del ciclo hidrológico: La descripción del ciclo hidrológico se resume en el estudio de las precipitaciones y la evapotranspiración, ya que en respuesta a la complejidad de la medida de las recargas de los acuíferos se toma la cuenca como impermeable.

Precipitación: De acuerdo con la definición clásica de la Hidrología, la precipitación son los procesos mediante los cuales la lluvia cae a la superficie terrestre, lo que incluye la lluvia, la nieve y otros procesos tales como granizo y nevisca [4]. En nuestro contexto, la precipitación se refiere preferiblemente a la lluvia como el tipo más común que se presenta, con algunas eventualidades en las que se dan lluvias con granizo. Esta variable atmosférica, que en gran parte describe el tiempo atmosférico en nuestro país, exhibe una alta variabilidad espacial y temporal propia de procesos altamente caóticos y no lineales. La ubicación geográfica de Colombia, en medio de los dos océanos, a una latitud cercana al Ecuador y con unas condiciones topográficas tan variables, hace que los patrones de precipitación sean difíciles de identificar y por lo tanto requieren un estudio detallado.

Evapotranspiración: “El concepto de evapotranspiración fue introducido por primera vez por Charles Thornthwaite (Thornthwaite, 1948), quien la definió como la máxima cantidad de agua que puede evaporarse desde un suelo completamente cubierto de vegetación, que se desarrolla en óptimas condiciones, y en el supuesto caso de no existir limitaciones en la disponibilidad de agua” [5].

Para medir la evapotranspiración se utilizan estaciones especialmente diseñadas para esta variable llamadas lisímetros (ver Figura 1), debido a la complejidad en la estructuración de estas estaciones se utilizan fórmulas que dependen de otras variables meteorológicas un poco más fáciles de medir. Algunas de estas fórmulas se pueden observar en la Tabla 1.

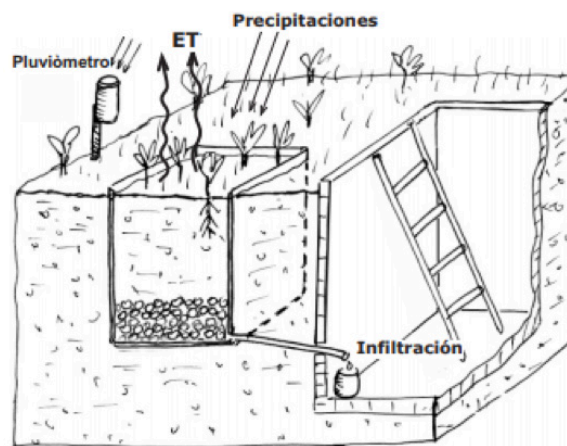


Fig. 1. Estación de medición de evapotranspiración (lisímetro)

Fuente: Tomado de [6]

Tabla 1. Ecuaciones de cálculo de evapotranspiración

Método	Medidas necesarias	Otros datos necesarios
Thornthwaite	Temperatura	De la latitud por una tabla se obtiene el número teórico de horas de sol.
Jensen-Heise	Temperaturas (medias, máximas y mínimas del mes más cálido), altitud y radiación solar.	Tablas de número teórico de horas de sol y la radiación solar se puede estimar.
Hargreaves	Temperatura y radiación solar	La radiación solar se puede estimar con temperaturas máximas y mínimas diarias.

Método	Medidas necesarias	Otros datos necesarios
Blanney-Criddle	Temperatura	Tablas de número teóricos de horas de sol y coeficiente que depende del tipo de cultivo que se encuentre en la cobertura vegetal.
Turc	Temperatura, y horas reales de sol	De las horas de sol se obtiene la radiación global incidente (cal/cm ² .día) con una fórmula.
Penman	Temperatura, horas reales de sol, velocidad del viento y humedad relativa	Por tablas se obtienen otros parámetros necesarios.

Fuente: Elaboración propia.

Temperatura: En meteorología, la temperatura se registra en las estaciones meteorológicas, de las que existen miles en todo el mundo. En estas estaciones se miden, por ejemplo, datos de temperatura a horas fijas, valores de temperaturas máximas y mínimas o se toman registros continuos en el tiempo, llamados termogramas.

Para analizar la distribución de la temperatura sobre grandes áreas se usan las isotermas, que son curvas dibujadas sobre un mapa que unen los puntos de igual temperatura. El cambio de temperatura en una dirección determinada del espacio, se llama gradiente de temperatura y se puede obtener del mapa de isotermas. Al analizar los gradientes de temperatura en los mapas, se puede deducir que donde las isotermas están más juntas, el cambio de temperatura en la región considerada es grande, es decir, el gradiente de temperatura es grande, y donde están más separadas, el cambio o gradiente es pequeño [7].

B. Redes meteorológicas

Una red de estaciones constituye un sistema articulado de medición, almacenamiento, envío y análisis de información relacionada con diferentes variables meteorológicas. Dicho sistema debe tener como objetivo fundamental el estudio de un fenómeno dado. En el caso que aquí atañe, se pretende recolectar y analizar información referida a las variables que con mayor claridad influyen en la estimación del ciclo hidrológico.

Componentes de una red de estaciones meteorológicas: La red meteorológica en su estructura considera elementos de medición, de almacenamiento de datos, de transmisión de datos y de interpretación de datos.

Estaciones: Los principales elementos que integran una red meteorológica son las estaciones. Una estación debe contener dispositivos que permitan medir, entre otras variables, presión barométrica, evapotranspiración, temperatura, humedad, precipitación, humedad del suelo, radiación solar, radiación ultravioleta, y velocidad y dirección del viento. Aparte de los sensores encargados de registrar una señal análoga referida a la medición de la variable en cuestión, las estaciones poseen una unidad central que transforma las señales análogas en digitales. Dicha unidad suele recibir el nombre de Unidad Inteligente de Procesamiento, y cuenta, además, con mecanismos sistematizados que permiten tomar decisiones en función de las condiciones actuales [8].

III. ÁREA DE ESTUDIO

El área de estudio seleccionada para este trabajo de investigación es la cuenca del río Negro en el Oriente antioqueño. El área de la cuenca se encuentra en jurisdicción de Cornare. Geográfica y topográficamente, el área de la cuenca del río Negro está localizada en la vertiente oriental de la Cordillera Central.

En la Figura 2, se presenta la delimitación de la cuenca [9], la cual es, por definición, la porción de terreno que drena al punto de cierre en la entrada del río Negro al embalse Peñol-Guatapé. La delimitación de la cuenca ha sido definida mediante las direcciones de flujo y la acumulación de flujo, en un proceso basado en el uso de un sistema de información geográfica, este procedimiento es definido en [10], a partir del DEM obtenido de estudios realizados por Cornare en la zona de estudio, donde luego de corregir vacíos se logró obtener los mapas de dirección de flujo, flujos acumulados y, por último, la delimitación de la cuenca [10].

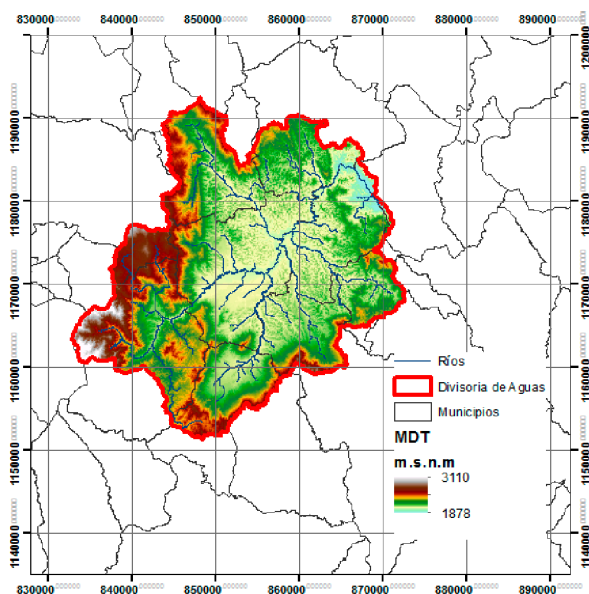


Fig. 2. Delimitación de la cuenca del río Negro

La cuenca del río Negro la conforman 219 veredas que se encuentran repartidas en diez municipios de la subregión del Oriente antioqueño: Rionegro, Marinilla, Santuario, San Vicente, El Retiro, El Peñol, Envigado, El Carmen de Viboral, Guarne y La Ceja del Tambo [9].

IV. METODOLOGÍA PARA EL PROCESAMIENTO DE LOS DATOS

Para la delimitación del área de estudio, se utilizaron los métodos y procesos que pueden

emplearse en la tarea de levantamiento y arreglo de los geodatos, y procesos de geomodelación ligados a la construcción de modelos espaciales con propósitos hidrológicos a partir de la cartografía disponible en estudios previos elaborados en el área de estudio por Cornare. Para el análisis de los datos espaciales se utilizó la herramienta de ESRI ArcGIS 10.6, con la cual se desplegaron los datos para las respectivas simulaciones y la utilización de herramientas específicas como Geostatistical Analyst and Spatial Analyst.

En el proceso de la delimitación de la cuenca, se toman las curvas de nivel encontradas en estudios realizados por Cornare para la cuenca del río Negro, donde se ejecutaron algunos ajustes, como la corrección en los vacíos de información resultantes en los Modelos Digitales de Superficie (*Digital Surface Model –DSM–*).

En este sentido, en [10] se propuso un método para la corrección de la cartografía, con el objetivo de lograr una representación actualizada de los elementos hidrológicos de superficie realmente significativos que permiten construir una condición de frontera. Por eso, para la delimitación del área de estudio se tomó el DSM de los estudios de Cornare y se utilizó el método de corrección de vacíos ya expuesto [11].

En busca de la actualización de la información física, se determinaron diferentes metodologías en función de la variable de interés, como delimitación de la cuenca, precipitación y evapotranspiración.

Análisis de precipitación: Los datos de precipitación se analizaron mediante procesos estadísticos (tendencia central y variabilidad) y técnicas geoestadísticas [12]. La calidad de la estimación de la precipitación está sujeta a la calidad, cantidad y distribución de los datos de precipitación disponibles en la zona de estudio. La incertidumbre asociada con la estimación va a aumentar en la medida en que la cantidad de datos de precipitación disponible sea baja.

Para determinar la variabilidad en las precipitaciones en el área de trabajo, se tomaron las diferentes estaciones de las cuales se tiene información (ver Tabla 2).

Tabla 2. Estaciones pluviométricas utilizadas en el análisis de precipitación

Nombre	Coordenadas		Cota	Fuente
	Latitud	Longitud		
Tulio Ospina	6,319694	-75,55325	1438	Ideam
Campoalegre	6,073778	-75,335611	1850	Ideam
La Unión	5,998333	-75,381389	2530	Ideam
Santuario	6,133056	-75,273583	1250	Ideam
La Salada	6,048361	-75,624667	1923	Ideam
A. J. M. Córdova	6,168667	-75,425917	2157	Ideam
La Cuchilla	6,368333	-75,453611	1600	Ideam
Corrientes	6,311944	-75,253556	1965	Ideam
El Peñol	6,214222	-75,241306	2047	Ideam
Montebello	5,950833	-75,536833	560	Ideam
La Selva Autom	6,131917	-75,414556	2125	Ideam
Sta. Helena	6,196889	-75,51675	2550	Ideam

Fuente: Elaboración propia.

La selección de las estaciones se realiza utilizando los polígonos de Thiessen, los cuales consisten en la delimitación de áreas de influencia a partir de un conjunto de puntos; el tamaño de éstos y la configuración dependen de la distribución de los puntos originales de las estaciones que se van a utilizar (ver Figura 3).

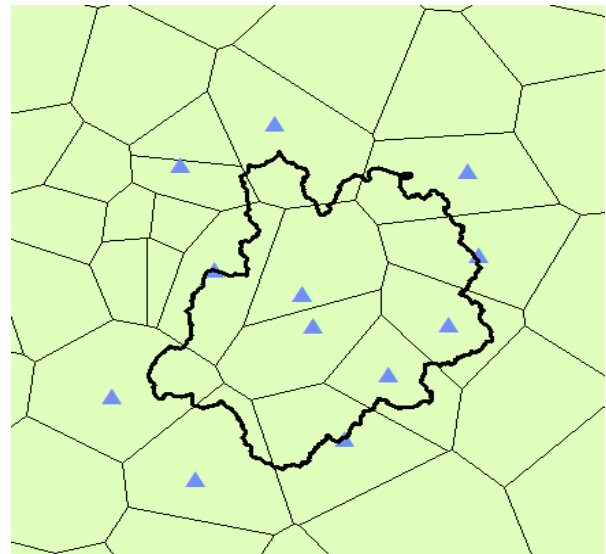


Fig. 3. Polígonos de Thiessen de interés para el área de trabajo

En total se tomaron los datos de 12 estaciones en un periodo de 20 años, comprendidos entre 1998 y 2017. En la Tabla 2 se presentan las estaciones utilizadas, así como su ubicación y la fuente de información.

Análisis de evapotranspiración: La evapotranspiración (ET) como variable hidrológica a nivel territorial, es primordial para cualquier tipo de análisis concerniente con el manejo del agua (gestión de cuencas hidrográficas, embalses locales, entre otras) [13].

Para el estudio de la ET, se tomaron los datos de tres estaciones suministradas por el Ideam las cuales se detallan en la Tabla 3. Luego de obtener los datos se realizó una interpolación de distancia inversa ponderada (*Inverse Distance Weighted – IDW*–).

Tabla 3. Estaciones de evapotranspiración

Nombre	Coordenadas		Cota	Fuente
	Latitud	Longitud		
La Salada	6,048361	-75,624667	1923	IDEAM
Corrientes	6,311944	-75,253556	1965	IDEAM
La Selva Autom	6,131917	-75,414556	2125	IDEAM

Fuente: Elaboración propia.

Luego de procesar la información se obtienen los mapas de precipitación y evapotranspiración anuales de la cuenca del río Negro, desde el año 1998 al 2017, para el caso de precipitación, y de 1994 al 2003, para la evapotranspiración, los cuales se visualizan en la sección de resultados.

Análisis estadístico de los datos obtenidos:

Los datos de las estaciones seleccionadas para el análisis de evapotranspiración y precipitación son descargados directamente de la página del portal de información del Ideam, en dichas series faltan algunos datos, lo cual se corrigió utilizando el método de Criterio de Valor Medio, que consiste en sustituir la carencia de un registro por el valor medio de la serie [14]. En la Tabla 4 se presentan los datos de precipitación ya completados para los años comprendidos entre 1998 y 2003.

Tabla 4. Datos de precipitación de 1998 a 2003

Nombre Estación	Precipitación Anual					
	1998	1999	2000	2001	2002	2003
Tulio Ospina [27015090]	4491,4	5720,6	4573,4	1962,8	2417,9	2263,4
Campoalegre [23080650]	2983,8	3408,0	3693,0	1905,9	2490,2	2465,6
La Unión [26180170]	2010,0	3085,6	2876,0	2008,5	1970,5	1560,5
Santuario [23080920]	2500,1	3147,9	3254,4	2580,1	2382,4	2711,8
La Salada [27015260]	2661,5	3418,6	3125,3	2065,0	2132,9	2592,2

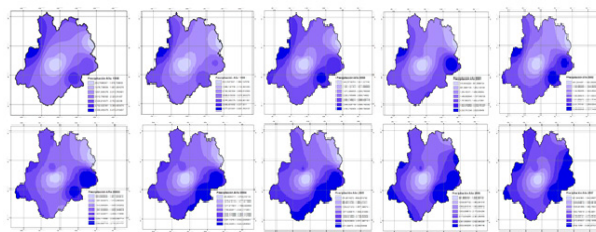
	Precipitación Anual					
A. J. M. Córdoba [23085200]	1928,3	2536,0	2470,8	1638,9	1594,6	1931,7
La Cuchilla [27010820]	2020,3	2230,9	2466,7	1367,9	1555,6	1928,2
Corrientes [23085160]	3150,0	2527,7	2646,7	943,9	931,6	1360,1
El Peñol [23085110]	406,0	702,5	797,5	562,8	605,1	484,8
Montebello [26180200]	2088,5	3481,6	2629,9	1834,1	2246,3	2551,9
La Selva Autom [23085260]	222,7	202,2	445,1	310,9	481,1	669,0

Fuente: Elaboración propia.

V. RESULTADOS

En el procesamiento de la información de precipitación se cuenta con una mayor cantidad de datos, lo que ofrece mayor confianza en los resultados obtenidos. En el análisis de evapotranspiración solo se cuenta con cuatro estaciones, lo que podría incurrir en algunos vacíos de información o en la reducción de su calidad. A continuación, se muestran los resultados para cada uno de los análisis realizados.

Precipitación: Después de haber hecho el análisis de las 12 estaciones pluviométricas y una interpolación de distancia inversa ponderada (IDW), se llega a los siguientes mapas de precipitaciones anuales en el área de estudio. Las unidades de medida están dadas en mm de lluvia.

**Fig 4.** Precipitación entre los años 1998-2007

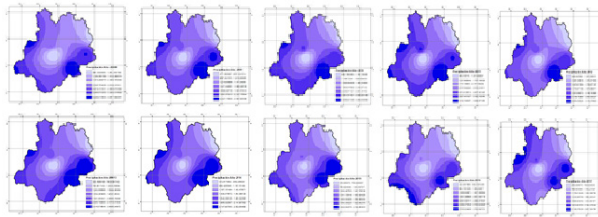


Fig. 5. Precipitación entre los años 2007-2017

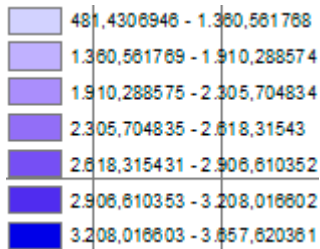


Fig. 6. Escala de variación de la precipitación

Evapotranspiración: Para la realización del análisis de evapotranspiración en la cuenca del río Negro se utilizan las tres estaciones con las que se cuenta. Después de realizada la interpolación de distancia inversa ponderada (IDW), se llega a los siguientes mapas de evapotranspiración anuales en el área de estudio. Las unidades de medida están dadas en mm.

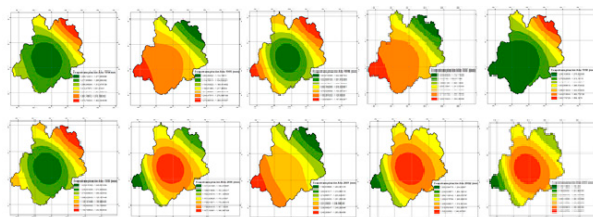


Fig. 7. Evapotranspiración entre los años 1994-2003

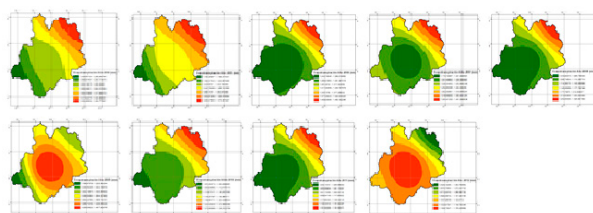


Fig. 8. Evapotranspiración entre los años 2003-2012

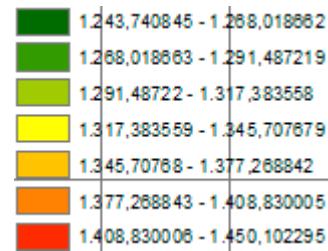


Fig. 9. Escala de variación de la evapotranspiración

Luego de realizar el análisis de las evapotranspiraciones desde el año 1994 hasta el año 2012, para un total de 19 años en la cuenca del río Negro, se llega a una evapotranspiración media la cual se puede visualizar en la Figura 8.

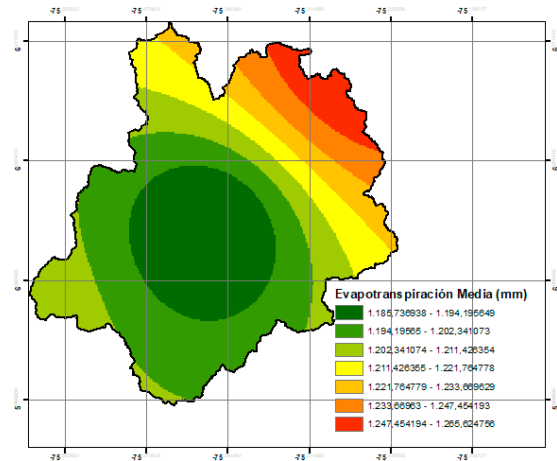


Fig. 10. Evapotranspiración media anual (mm)

Los valores de la evapotranspiración media oscilan entre 1185,73 mm y 1265,42 mm anuales.

Temperatura: Al momento de realizar el análisis de temperaturas para la cuenca del río Negro se llega a un obstáculo: la falta de estaciones de temperatura en el área de la cuenca. Por tal motivo se procedió a utilizar el método expuesto en [15], el cual aporta las fórmulas de temperatura media, máxima y mínima para las diferentes regiones de Colombia, dependiendo de la altura con respecto al mar. En el caso de la cuenca del río Negro, se utilizan las propuestas para la región Andina, donde se tienen las siguientes ecuaciones:

$$T_m = 29,42 - 0,0061 A$$

$$T_{max} = 33,97 - 0,0058 A$$

$$T_{min} = 23,97 - 0,0059 A$$

Donde:

T_m : Temperatura media

T_{max} : Temperatura máxima

T_{min} : Temperatura mínima

A: Altura con respecto al nivel del mar

Para la utilización de estas fórmulas se emplean las curvas de nivel, seguidamente, por medio del software ArcGis, se hace el cálculo de temperatura media, máxima y mínima con cada una de las curvas de nivel, para luego realizar una interpolación triangular (*Triangulated Interpolation -TIN-*) y así poder obtener un modelo del comportamiento de la temperatura en la cuenca. En las Figuras 9, 10 y 11 se puede visualizar cómo es el debido comportamiento de las temperaturas en la cuenca [16].

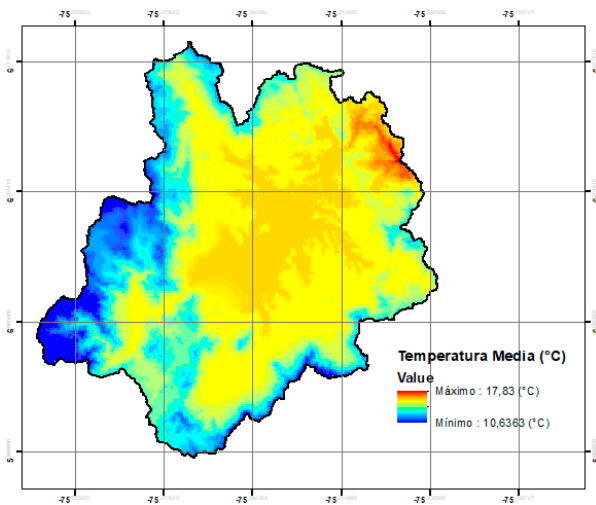


Fig. 11. Mapa de temperatura media (°C)

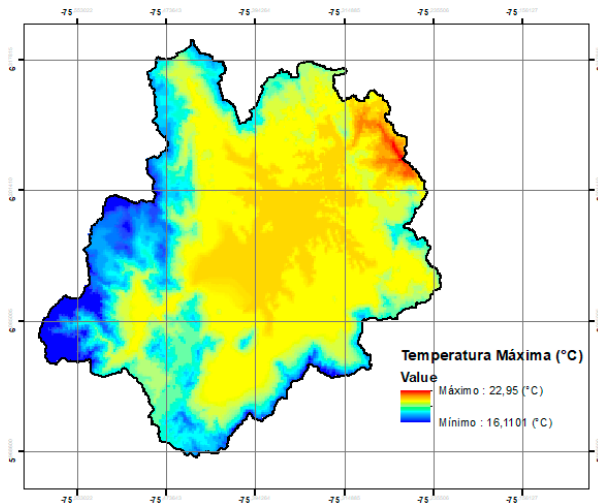


Fig. 12. Mapa de temperatura máxima (°C).

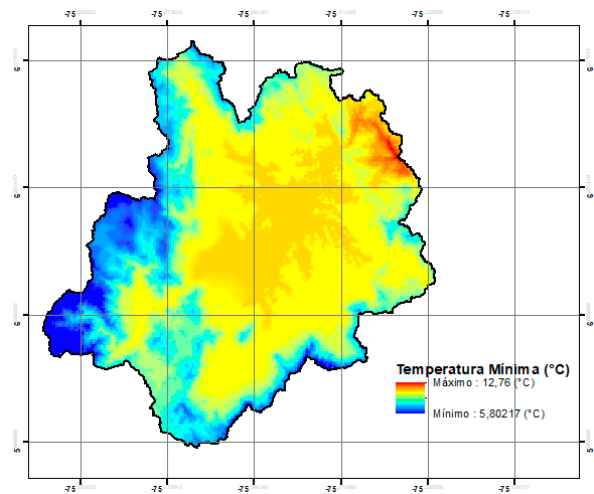


Fig. 13. Mapa de temperatura mínima (°C).

VI. DISCUSIÓN

Luego de obtener los valores teóricos y prácticos (por medio de las estaciones) de la evapotranspiración, estos pueden ser comparados con los valores de referencia obtenidos por [17].

La caracterización física de la cuenca del río Negro proveerá información para mejorar la fase de entrenamiento de la red neuronal que se utilice para la construcción del modelo predictivo de nivel del río Negro.

Para llevar a cabo las calibraciones o correcciones de los datos de evapotranspiración, se calcula el caudal del río Negro haciendo uso de los mapas de precipitación y evapotranspiración por el método del balance a largo plazo, y luego se compara con el caudal medio multianual extraído de los datos de una estación de caudal que se encuentre sobre el río Negro.

VII. CONCLUSIONES

La cantidad de estaciones que se cuentan para el análisis de la precipitación ofrece una buena distribución de los datos dentro de la cuenca del río Negro. La caracterización física de la cuenca permitirá realizar un mejor entrenamiento del modelo predictivo de nivel del río Negro.

Para determinar la calidad de la información procesada con respecto a la evapotranspiración, es necesario realizar una comparación con métodos basados en datos de temperatura y altitud, debido a la poca cantidad de estaciones con las que se cuenta para esta variable.

REFERENCIAS

- [1] A. Viera Brones, A., López Santos, J. L. González Barrios, & R. Trucios Caciono, “Caracterización de la cuenca Alip con la ayuda de un sistema de información geográfica”, *Revista Chapingo Serie Aridas*, XI, 23-34, 2012.
- [2] F. Amiri, “Estimate of erosion and sedimentation in semi-arid basin using empirical models of erosion potential within a Geographic Information System”, *Air, Soil and Water Research*, 2010. <https://doi.org/10.4137/ASWR.S3427>
- [3] P. Llorens, “La evaluación y modelización del balance hidrológico a escala de cuenca”, *Ecosistemas*, 1, 2003. Disponible en: <http://www.aeet.org/ecosistemas/031/opinion1.htm>
- [4] V. T. Chow, D. R. Maidment, y L. W. Mays, *Hidrología aplicada*, Bogotá: McGraw Hill, 1994.
- [5] J. A. Gómez y M. C. Cadena, “Validación de las fórmulas de evapotranspiración de referencia (ET_o) para Colombia”, Instituto de Hidrología, Meteorología y Estudios Ambientales (Ideam), Subdirección de Meteorología, 2017.
- [6] F. J. Sánchez San Román, “Evapotranspiración”, Departamento de Geología, Universidad de Salamanca, 2000. Disponible en: <http://hidrologia.usal.es/temas/Evapotransp.pdf>.
- [7] J. Inzunza, *Meteorología Descriptiva y aplicaciones en Chile*, Universidad de Concepción, 2003. Disponible en: http://nimbus.com.uy/weather/Cursos/Curso_2006/Textos%20complementarios/Meteorologia%20descriptiva_Inzunza/cap1_Inzunza_La%20atmosfera.pdf
- [8] G. A. Martínez López, “Meteorological Station for forming networks”, *Installation Process*, vol. 21, N.º 1, pp. 115-121, 2016.
- [9] Consorcio POMCAS Oriente Antioqueño, Formulación del plan de ordenación y manejo de la cuenca hidrográfica de río Negro. Documento introductorio del POMCA río Negro, 2016.
- [10] E. Pérez, C. Castro, Y. Castrillón, G. Valencia y J. Valdés, Comps., *Geoinformática Aplicada con aprendizaje basado en problemas*. Medellín: Instituto Antioqueño de Investigación, 2017.
- [11] Ó. Quintero y G. Valencia, “Corrección de vacíos de información (voids) a partir de la aplicación de promedios de diferencias de altura entre superficies de elevación”, en *Geoinformática Aplicada con aprendizaje basado en problemas*, E. Pérez, C. Castro, Y. Castrillón, G. Valencia y J. Valdés, Comps. Medellín: Instituto Antioqueño de Investigación, 2017, pp. 62-77.

- [12] R. Giraldo Henao, *Introducción a la geoestadística: Teoría y aplicación*, Bogotá: Universidad Nacional de Colombia, 2011.
- [13] P. J. Pérez y F. Castellví, “Análisis de la evapotranspiración a escala local y regional en Cataluña”, *Ingeniería del Agua*, vol. 9, pp. 59-72, 2002.
- [14] R. Rodríguez, M. C. Llasat y J. Martín-Vide, *Análisis de Series Temporales en Climatología: modelización y homogeneidad*. Barcelona: Ediciones de la Universitat de Barcelona, Textos Docentes N.º 165, 2000.
- [15] B. Chaves Córdoba y A. Jaramillo Robledo, “Regionalización de la temperatura del aire en Colombia”, *Cenicafé*, N.º 49, 224-230, 1998.
- [16] J. F. Escobar, T. Betancur y J. C. Valdés, “Herramientas SIG para la reconstrucción paleotopográfica de un ambiente sedimentario”, *Geotecnologías, Herramientas para la construcción de una nueva visión del cambio global y su transformación para un futuro sostenible*, N.º XVII, pp. 2241-2254, 2016.
- [17] A. Jaramillo, “Evapotranspiración de referencia en la región Andina de Colombia”, *Cenicafé*, vol. 57, N.º 4, pp. 288-298, 2006.

Capítulo XI

Plataformas de Sensores Remotos en la Generación de Cartografía como Instrumento de Apoyo a la Gestión de Ecosistemas Estratégicos

John Escobar - john.escobar@udea.edu.co

Alexandra Urán - alexandra.uran@udea.edu.co

Universidad de Antioquia, , Medellín, Colombia

Vanesa Paredes - contactenos@corpouraba.gov.co

Corpouraba

Juan Carlos Valdés - jcvaldes@elpoli.edu.co

Juan Parra - jcparra@elpoli.edu.co

Esteban Zapata - eszatre@gmail.com

Politécnico Colombiano Jaime Isaza Cadavid, , Medellín, Colombia

I. INTRODUCCIÓN

En Colombia, los ecosistemas estratégicos han adquirido una significativa importancia en los últimos años, no sólo en términos de investigación científica, sino también de cambios jurídicos y políticos en su manejo. Es así como desde inicios de la década de 2010 se han adelantado estudios interdisciplinarios sobre ecosistemas de páramo, con el fin de caracterizarlos ecológica, geográfica y socioculturalmente. Producto de estos esfuerzos, en el año 2016 se consolidó la política de delimitación de páramos, la cual se convirtió en un elemento determinante para la legislación sobre explotación minera y producción agropecuaria en áreas de páramo. Los resultados se consignan en una serie de resoluciones, entre las que se resalta la Resolución 0496 del 22 de marzo de 2016, que delimita el complejo “Páramos del Sol – Las Alegrías”, con una extensión de 15.396 hectáreas, en jurisdicción de 11 municipios ubicados en los departamentos de Antioquia y Chocó, Colombia [1]. Los ecosistemas de páramos cobran gran relevancia por ser sistemas naturales singulares en cuanto a características biológicas, climáticas y de suelo [2]. Sin embargo, investigaciones recientes los han identificado como sectores con una alta vulnerabilidad por los efectos del cambio climático [3], el avance de la frontera agrícola [4] y la minería [5].

En enero de 2018, el grupo de investigación Medio Ambiente y Sociedad (MASO), acompañado por el Grupo de Ingeniería y Gestión Ambiental (GIGA), elaboraron una propuesta para desarrollar el proyecto “Construcción de una Herramienta de Planificación, Ordenamiento y Manejo del Complejo de Páramos Frontino-Urrao, Antioquia”, mediante convenio entre la Universidad de Antioquia y la Corporación para el Desarrollo Sostenible del Urabá (Corpourabá), labor en la cual se desarrolla una batería de instrumentos en los que convergen las redes de Actores, Factores y Sectores (FAS), las metodologías de Presión, Estado, Respuesta (PER) y la cartografía de actores sociales (MAC), apoyados en los sistemas de información geográfica (SIG) y los sensores remotos, como un aporte significativo a la gestión sostenible de estos ecosistemas y cuyos avances se presentan a continuación.

II. MATERIALES Y MÉTODOS

Uno de los enfoques de la sostenibilidad, denominada fuerte, plantea que esta debe preocuparse directamente por la salud de los ecosistemas en los que se inserta la vida y la economía de los seres humanos sin ignorar la incidencia que sobre los procesos del mundo físico tiene el razonamiento monetario [6]. Enmarcado en este concepto se encuentra el Capital Natural

Crítico (CNC), el cual puede entenderse como el conjunto de recursos medioambientales que, a una escala geográfica dada, asegura las funciones medioambientales importantes para las cuales no existe ningún sustituto en términos de capital manufacturado o humano [7]. Pero la instrumentación de este concepto conduce a la siguiente pregunta: ¿cómo identificar y evaluar estas funciones en una región y a una escala determinada? En este sentido, en [8] se propone un modelo para la identificación del CNC con apoyo en herramientas de manejo de información y análisis geoespacial. Escobar y Lopera [9] retoman esta propuesta y sugieren una formulación soportada en modelos presión-estado-respuesta (PER) espacialmente explícitos, elaborados con el apoyo de los sistemas de información geográfica (SIG) e información proveniente de cartografía temática y sensores remotos (ver Figura 1).

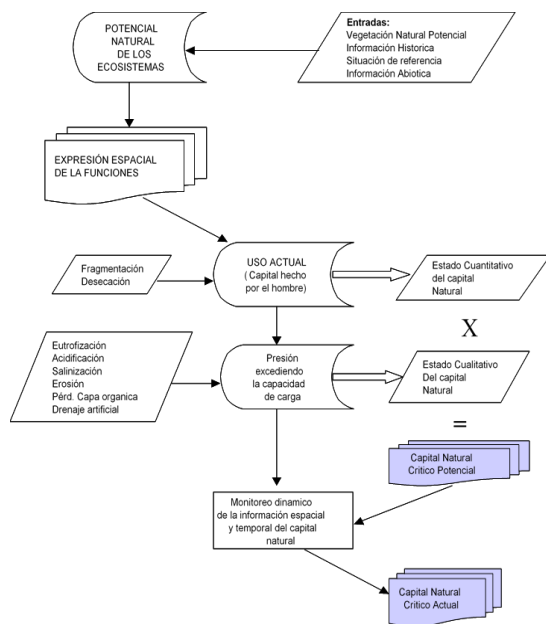


Fig. 1. Modelo conceptual propuesto por Van der Perk y de Groot para la identificación del CNC apoyado en SIG.

Fuente: Adaptado de [9].

A pesar de que esta propuesta permite aumentar la proximidad a las variables espaciales explícitas, todavía subyace un problema habitual

en los estudios territoriales, que consiste en cómo establecer de manera participativa unos escenarios futuros de gestión-interacción. Una respuesta aproximada a esta inquietud puede evidenciarse en [10], cuyos autores usan los SIG en la valoración de las funciones ambientales asociadas a los cambios de uso y degradación del ambiente natural, para analizarlos a través de funciones integradoras. Así, brindan un aporte adicional al abordar el problema de la escala en tres dimensiones: la espacial, la temporal y la social, las cuales están estrechamente relacionadas a preguntas espacialmente explícitas. En este sentido, proponen una metodología consistente en: 1) Determinación del confinamiento del estudio (área de estudio); 2) identificación de aspectos que influyen en la identidad del territorio; 3) el análisis de estos aspectos y su influencia en la calidad y el manejo del territorio; 4) la exploración de los efectos de los cambios de uso sobre la identidad del territorio, y 5) la investigación de la existencia de percepciones asociadas al uso del paisaje. Finalmente, proponen un sistema combinado de modelos, métodos e indicadores que pueden ser usados en la definición de políticas de usos futuros soportados en la identificación de Factores-Actores y Sectores (ver Tabla 1).

Adicionalmente, Risler y Ares [11] sostienen que el ejercicio de mapear es, en esencia, una representación del espacio que alberga subjetividades como formas de interpretación del mundo. Cuando definimos un territorio, más allá de los elementos físicos y geomorfológicos que lo modelan, entendemos que existen relaciones políticas, sociales y culturales entre diversos actores que, de acuerdo con sus intereses, configuran el entorno a partir de distintas perspectivas del territorio. La cartografía convencional contiene simbologías y convenciones que no suelen identificarse con la construcción territorial de las comunidades, y es allí donde el ejercicio de la cartografía social cumple un papel importante como forma de representación por fuera de los límites de esa cartografía, plasmando la complejidad de la estructura social y sus procesos de territorialización que rebaten las narraciones construidas desde discursos hegemónicos.

Tabla 1. Sistema combinado de modelos, métodos e indicadores para definición de políticas de usos futuros del suelo

	Paso	Estrategias y practicas	Tipo de Información	Método	Entrada de Información
Preguntas Clave 2	1.1. Identifique las cuestiones claves y centre en las percepciones del tomador de decisiones	Dialogo entre científicos y tomadores de decisiones	Cualitativa	Talleres, seminarios de trabajo	
	1.2 Lista de Indicadores	Retroalimentación entre científicos y tomadores de decisiones	Tan cuantitativa como sea posible	Redes de Impactos/Modelos	1.1.
	1.3 Definición de Factores, Actores y Sectores (FAS) relevantes	Retroalimentación de científicos	Cualitativa	Redes FAS	1.1 / 1.2
	1.4 Selección casos de estudio	Actividades científicas	Mapas con zonas socioeconómicas y biofísicas	Modelos SIG preliminares	1.2 / 1.3
Preguntas Clave 2	2.1 Antecedentes históricos	Actividades científicas	Cualitativa	Escenarios de las redes FAS existentes	1.3
	2.2 Historia de los casos de estudio seleccionados	Información de representantes locales	Cualitativa	Talleres de trabajo con representantes locales	2.1
	2.3 Proceso de toma de decisiones	Información de representantes locales y regionales	Cualitativa	Juego de roles	
	2.4 Implicaciones en los Factores, Actores y Sectores y los indicadores	Actividades científicas	Cuantitativa	Redes FAS	2.1 / 2.2
Preguntas Clave 3	Preguntas Clave 3	Actividades científicas	Cuantitativa	Modelos exploratorios Integrados (multidominio, multiescalar y multifactores)	1.3 / 2.1
	3.1 Modelos de tendencias a gran escala	Actividades científicas	Cuantitativa	Modelos basados en agentes y de cambios de coberturas y usos	1.4 / 3.1
	3.2 Modelos de comportamiento de agentes en los casos de estudio	Actividades científicas	Semicuantitativa	Sistemas de Soporte a las decisiones	3.1 / 3.2
	3.3 Interface políticos - Científicos	dialogo	Semicuantitativa	Talleres, trabajo con grupos objetivo	3.2 / 3.3

Fuente: Tomado de [10].

Conscientes de que la existencia de un instrumento por sí mismo no garantiza que las variables implícitas evidenciadas por los colectivos sociales puedan ser expresadas en términos espaciales, máxime cuando se deben cumplir con reglas cartográficas y geográficas rigurosas, se procedió a retomar algunos instrumentos técnicos robustos, a saber: la cartografía básica y temática desarrollada por los entes territoriales con jurisdicción en la zona, las imágenes provenientes de los sensores satelitales Landsat 8 y Sentinel 2, y las imágenes topográficas SRTM corregidas y liberadas en fechas recientes para la región de estudio, todo ello combinado con impresos y escenificaciones elaborados con el SIG-ArcGis, que permitieron desarrollar talleres adaptando metodologías provenientes de la cartografía social, para identificar procesos sobre el territorio que solo pudieron ser evidenciados gracias a la interacción directa con las comunidades.

III. RESULTADOS Y DISCUSIÓN

Como punto de partida, se plantea que una metodología de construcción colectiva basada en las redes FAS puede contribuir a resignificar elementos de concertación que, de otro modo, pueden diluirse dentro de la problemática en general. Igualmente, se reinterpretan algunos resultados obtenidos en la zonificación de hechos de la naturaleza, logrados a partir de la cartografía convencional, enriquecida por conceptos y técnicas basados en la evaluación de funciones y servicios ecosistémicos, análisis de sensores remotos y evaluaciones de potenciales hidrológicos.

La región de estudio abarca 26 veredas¹ de los municipios de Urao, Abriaquí, Cañasgordas, Frontino y Giraldo, pertenecientes a la jurisdicción de Corpourabá, en las cuales se identifican ocho sectores que representan los páramos de la región. En ellos sobresale el denominado El Sol - Las Alegrías, que por sí solo constituye el 66 % del área declarada en la Resolución 0496 de 2016. En la Figura 2 se muestra la localización de este complejo.

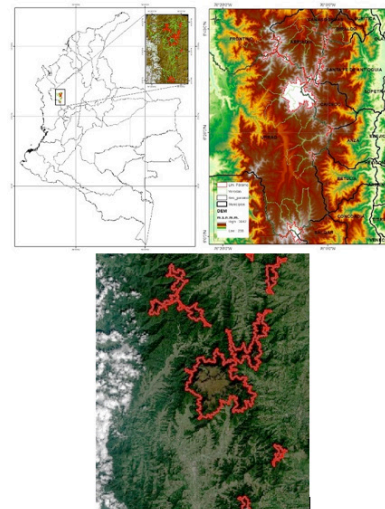


Fig. 2. Localización del Complejo de Páramos Frontino-Urao (CPFU) con detalle en una composición Landsat Color Real (derecha) sobre el Corredor de las Alegrías-El Sol (nótese el color marrón asociado a la vegetación propia del páramo)

1 División administrativa subordinada a la municipalidad.

Las actividades desarrolladas, a pesar de no haberse ejecutado en estricto orden cronológico, se pueden agrupar a través de las siguientes acciones:

Talleres de cartografía: aproximándonos a un lenguaje común. Con algunos productos cartográficos (analógicos), proyecciones y animaciones se realizó una primera jornada de talleres con una audiencia tan plural como fue posible convocar. Así, bajo la orientación de un animador, se les pidió hacer la ubicación más precisa posible de adhesivos que ilustraban problemáticas (como explotación ilegal de madera), eventos de incidencia territorial notable (como cultivos extensos de frutas), potencialidades (como paisajes escénicamente memorables), y acciones que desde la sociedad civil propenden por la protección del ecosistema (como observatorios regulados de aves). De esta forma, y en cerca de una decena de talleres realizados en diferentes veredas, se logró identificar hechos que se pueden denominar “eventos geográficos”, dado que su localización nos permite sistematizarlos por asociación con elementos geográficos explícitos, como son la vereda o incluso la microcuenca (ver Figura 3).



Fig. 3. Talleres participativos de localización de eventos geográficos. Derecha: resultado de uno de los grupos en la vereda El Chuscal.

Identificando Factores, Actores y Sectores. Este ejercicio, también evidentemente participativo, consistió en la identificación de actores y su respectivo posicionamiento social. Este método de campo es también conocido como mapas sociales o

sociogramas [12], pues el mapeo de actores claves (MAC) supone el uso de esquemas para representar la realidad social y establecer estrategias de cambio, que, en nuestro caso, tienen que ver con una realidad socioambiental que se deconstruye hacia una estrategia de manejo integral de los recursos del páramo y sus áreas de influencia. Debe aclararse que no solo se trata de un listado de posibles actores presentes en el territorio, sino que también se busca conocer sus acciones y los objetivos de su presencia y sus perspectivas de futuro.

Se parte de una lista exhaustiva de factores ambientales geográficamente reconocibles, y a partir de allí se pide a los participantes que identifiquen los distintos actores con presencia en la zona del páramo y posibles áreas de amortiguamiento. Los actores fueron previamente categorizados por el equipo de investigadores en cuatro grupos básicos: sector público, sector privado, organizaciones de la sociedad civil, y una última categoría de gestores de conocimiento (públicos, privados y Organizaciones No Gubernamentales —ONG—). Se propone su ubicación en un plano cartesiano que confronta su presencia y relevancia en la toma de decisiones de acciones y actividades en el páramo.

La Figura 4 ilustra el resultado de uno de estos talleres. De forma anecdótica, puede evidenciarse que algunas instituciones presentan una relevancia alta, pero presencia baja, aunque la calificación asignada a la relevancia se otorga más en el sentido de la “necesidad”. Por otro lado, se evidencian amontonamientos con especial énfasis en los factores que emergen como amenazas o sectores asociados al uso de los recursos naturales, lo cual en general parece ser una tendencia rechazada para el páramo y sus zonas aledañas.

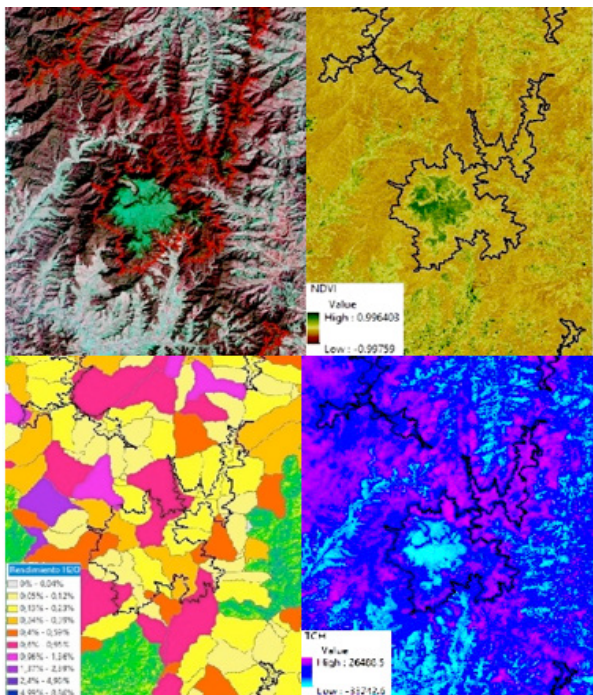


Fig. 5. Detalle sobre el páramo de El Sol-Las Alegrías, de algunos insumos obtenidos a partir de sensores remotos.

Nótese el color esmeralda rodeado de marrón (páramo y bosques circundantes) en el falso color (izquierda), y las respuestas disímiles del índice de vegetación (alto valor) y del Tasseled Cap Humedad (bajo valor), mientras que el modelo de acumulación de flujos muestra valores que dependen predominantemente de la pendiente y el aspecto de ella.

Dado que se detecta una enorme variabilidad en los insumos analizados, y que algunos fenómenos solo son perceptibles a escalas detalladas, se privilegiaron los resultados que se podían obtener de las imágenes Sentinel 2 y de la cartografía oficial escala 1:25.000. Fue así que se establecieron, para estos productos y escalas, unas ponderaciones relativas del estado de las cuatro características arriba mencionadas, al igual que la presión que ellas mismas podrían sufrir. La Figura 6 ilustra estas valoraciones.

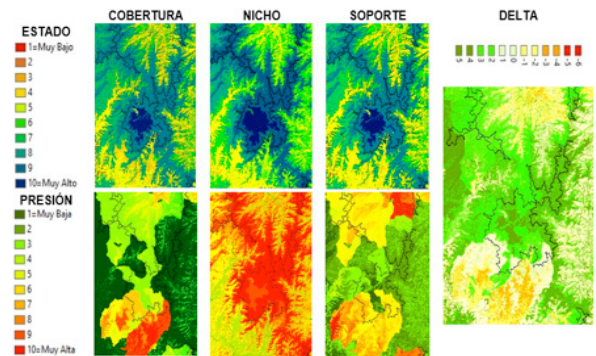


Fig. 6. Valoración espacial de las variables de Estado y Presión apoyadas en imágenes Sentinel 2 en las bandas de celda 10x10 m. Derecha: Ilustración espacial de la Presión vs. Estado si todas las variables tuviesen la misma ponderación.

IX. CONCLUSIONES

La fusión de metodologías cuantitativas y cualitativas puede involucrar múltiples inconvenientes y aproximaciones, sin embargo, en el trabajo expuesto es posible evidenciar que las metodologías participativas se pueden alimentar con información que se adquiere o procesa de forma muy técnica, pero, habitualmente, sin participación de las comunidades. En este caso, las imágenes Sentinel 2 fueron de gran ayuda para obtener una lectura actualizada del fenómeno estudiado. Ellas permitieron jerarquizar los estados de la naturaleza, al tiempo que la participación de las comunidades y la investigación social brindaron los insumos para ponderar y especializar las presiones. Aventurando un resultado final, es posible suponer que existirá un mapa delta que alerte a los gestores territoriales y habitantes de donde se presentan desequilibrios que puedan indicar la necesidad de proteger y recuperar, y donde este frágil equilibrio se mantiene, pero es aconsejable conservar, o aprender de él para usar este conocimiento en otras regiones y circunstancias.

Es importante resaltar que la utilización de información producto de sensores remotos, tales como Sentinel, Landsat, y modelos de elevación de radar como Alos Palsar, posibilitan acceder de forma libre y rápida a información de superficies

terrestres de alta escala y fundamental para la modelación de escenarios para la gestión del territorio, la cual es difícil y costosa de levantar en tierra. Esta información satelital puede ser validada por instrumentos en tierra y por la cartografía social obtenida en los talleres con las comunidades presentes en las zonas de páramos.

REFERENCIAS

- [1] Ministerio de Ambiente y Desarrollo Sostenible de Colombia (Minambiente), Resolución 0496 del 22 de marzo 2016. “Por medio de la cual se delimita el Complejo de Páramos Frontino – Urrao “Páramo del Sol – Las Alegrías” y se adoptan otras determinaciones.
- [2] M. Morales, J. Otero, T. Van der Hammen, A. Torres, C. Cadena, C. Pedraza, N. Rodríguez, C. Franco, J. C. Betancourth, E. Olaya, E. Posada y L. Cárdenas, Atlas de páramos de Colombia, Bogotá: Instituto Alexander von Humboldt, 2007.
- [3] Grupo Intergubernamental de Expertos sobre el Cambio Climático (IPCC), Cambio climático 2007: Informe de síntesis, Ginebra, Suiza: OMM, PNUMA, 2007.
- [4] L. H. Estupiñán, J. Gómez, V. Barrantes y L. Limas, “Efecto de actividades agropecuarias en las características del suelo en el páramo El Granizo, (Cundinamarca, Colombia)”, Revista U.D.C.A Actualidad & Divulgación Científica, vol. 12, N.º 2, pp. 79-89, 2009.
- [5] E. Guerrero, Implicaciones de la minería en los páramos de Colombia, Ecuador y Perú. Documento de trabajo. Biblioteca virtual RS, 2009. [En línea]. Disponible en: <http://bibliotecavirtual.info/2011/04/implicaciones-de-la-mineria-en-los-paramos-de-colombia-ecuador-y-peru-documento-de-trabajo/> [Accedido el 23 de noviembre de 2013].
- [6] J. M. Naredo, “Sobre el origen, el uso y el contenido del término sostenible”, Documentación Social, N.º 102, pp. 129-147. Ed. Caritas Española.
- [7] J. F. Noël, y M. O’connor, “Strong Sustainability and Critical Natural Capital”. En Valuation for Sustainable Development: Methods and Policy Indicators. Faucheux, S., O’Connor, M., (Eds.). Edward Elgar Publisher, Cheltenham, 1998, pp. 75–99.
- [8] J. Van der Perk and R. de Groot, Towards a method to estimate Critical Natural Capital. An inventory of methods to determine critical natural capital in the Netherlands. Discussion paper for second meeting of the CRITINC-project, 30/11 – 1/12, 1998, Paris, France, 2000.
- [9] J. Escobar y S. Lopera, “El Concepto de Capital Natural Crítico Aplicado a la Gestión de Recursos Energéticos”, en Memorias del cuarto Taller Internacional de Energía y Medio Ambiente. Cienfuegos, Cuba: Centro de Estudios de Energía y Medio Ambiente (CEEMA), 2006.
- [10] K. Knickel and K. Kok, “Future Land Use in Europe: Trends, Challenges and Policy (FLU-E)”, International Workshop Transition in Agriculture and Future Land Use Patterns, December 1-3, 2003 - Wageningen, The Netherlands: Institute for Rural Development Research (IfLS) at J. W. Goethe – University Frankfurt (D).
- [11] J. Risler, y P. Ares, Manual de mapeo colectivo: recursos cartográficos críticos para procesos territoriales de creación colaborativa. Buenos Aires: Tinta Limón, 2013.
- [12] M. M. Ceballos, “Manual para el desarrollo del mapeo de actores claves –MAC. Elaborado en el marco de la consultoría técnica GITEC-SERCITEC, 2004.
- [13] EC-FAO, “Stakeholders Analysis”, Annex I to lesson “understanding the Users’ Information Needs”, Food Security Information for Action Programme, FAO-EU, 2006.
- [14] D. Maidment, ed., Arc Hydro: GIS for Water Resources, Redlands, CA, USA: ESRI, 2002.

Capítulo XII

Modelamiento de la Concentración de Contaminantes en Suelos por el Método de Diferencias Finitas Utilizando Esquemas de Primero y Segundo Orden

Wilson Alberto Contreras - wcontre@unipamplona.edu.co

Docente -investigador, Universidad de Pamplona, Colombia

José Orlando Maldonado, orlmaldonado@unipamplona.edu.co

Docente-investigador, Universidad de Pamplona, Colombia

John Alexander Bohada, jbohada@jdc.edu.co

Docente-investigador, Fundación Universitaria Juan de Castellanos, Colombia

I. INTRODUCCIÓN

Hasta la aparición de efectos ambientales adversos, el enfoque en el análisis de las propiedades de los pesticidas se concentraba en su efectividad para el control de plagas. Actualmente, es mucho más importante analizar su destino ambiental: toxicidad, transporte, degradación, etc. De este modo, para que un pesticida se pueda aplicar, la legislación exige, entre otros requisitos, la realización de una serie de ensayos de campo y de laboratorio que permitan evaluar su comportamiento en el medio natural. En concreto, se examinan:

- La toxicidad del pesticida y sus productos de degradación sobre plantas y animales.
- El destino químico y el transporte del pesticida en el suelo, el aire y el agua.

A partir de datos obtenidos experimentalmente se pueden elaborar modelos de simulación que, una vez calibrados, permiten ampliar el conocimiento sobre el sistema estudiado en situaciones que no se han medido de forma experimental. El modelado matemático como herramienta de descripción es aplicable en la mayoría de las áreas de conocimiento y permite el estudio del comportamiento de sistemas complejos.

Los modelos numéricos para el estudio del comportamiento de los contaminantes en suelos continúan siendo de interés en la comunidad académica y científica. En [1] se propone el uso de barreras capilares de dos y tres capas como una solución eficiente para minimizar la migración en el suelo y se desarrollan los métodos numéricos para su simulación, sin dar detalles de los métodos desarrollados. En [2] se desarrolla un modelo utilizando métodos de elementos finitos, para predecir el proceso de contaminación de suelos saturados de forma variable por los contaminantes transportados en la escorrentía. En [3] se han desarrollado experimentos de columna aplicados a diez suelos con propiedades contrastantes y representativos de los tipos de suelo de captación y usos de la tierra, con el fin de evaluar el riesgo de contaminación de los recursos hídricos en el contexto del Altiplano boliviano. Sin describir el modelo, se presentan los resultados obtenidos a partir de la simulación mediante elementos finitos. Los resultados permiten describir los riesgos potenciales de contaminación de las aguas superficiales y subterráneas.

Dada la importancia del estudio de los pesticidas, se ha considerado conveniente desarrollar un modelo sencillo del transporte de pesticidas en el suelo que pueda ser utilizado en clases prácticas de simulación. El uso de un modelo de este tipo permite

analizar el efecto de la aplicación de insecticidas y herbicidas. Para utilizar el modelo, hay que conocer las propiedades físico-químicas del producto estudiado. Estas características están recogidas en las etiquetas de los productos comerciales.

El modelo presentado se desarrolla utilizando como lenguaje de modelado Python, junto a las librerías Numpy y Matplotlib para desplegar los resultados gráficos.

II. DESCRIPCIÓN DEL MODELO

Se presenta un modelo para describir la lixiviación y la degradación de un pesticida en un tipo específico de experimento de la columna, para lo cual se estudiará la solución analítica dada por Freijer [4]. En dicho experimento de la columna [5], se prepara una mezcla de suelo y pesticida con una cierta concentración total inicial. Esta mezcla se coloca como una capa con un espesor de 2 cm encima de una columna del mismo tipo de suelo. La columna de tierra debe tener una longitud de 28 cm y un diámetro de 5 cm. La capa delgada es así bastante pequeña comparada con la longitud de la columna. El suelo en la columna está saturado con agua antes de colocar la capa en ella. El agua se suministra por encima, con una tasa conocida constante, que induce el transporte del pesticida de la capa superior que contiene el pesticida hacia abajo. El BBA [6] recomienda un período de lixiviación de dos días en un flujo de Darcy de . Alternativamente, US-EPA [7] sugiere, para residuos antiguos de pesticidas, un período de lixiviación de 45 días en un flujo de Darcy de . Aunque el último guion esté pensado para residuos antiguos, puede ser aplicado también a mezclas frescas de suelo y pesticida. En el fondo de la columna hay desagüe. En el suelo empacado y la capa delgada, el pesticida es redistribuido sobre las fases sólidas y líquidas, y degradado por microorganismos, lo cual hace que la concentración cambie en el tiempo y el espacio. Durante el experimento, el efluente es reunido y es analizado. Después de un cierto período en el que ha habido lixiviación, el experimento se detiene y la columna de suelo se secciona en varias capas, las cuales son analizadas de forma separada.

A. Desarrollo del modelo

Los cuatro procesos considerados cuando se modela la lixiviación en la columna son: advección y dispersión unidimensionales a través de la fase líquida, sorción a la fase sólida, y a la degradación biológica. La ecuación que describe el balance de masa en la columna es (las ecuaciones para este modelo son dadas por [4]):

$$\frac{\partial X}{\partial t} = -\frac{\partial J_s}{\partial z} - kX, \quad (1.1)$$

Donde X es la concentración total del pesticida, t es el tiempo, z es la distancia, y k es la constante de degradación. Esta tasa constante está relacionada con la vida media del compuesto mediante la ecuación (1.2):

$$T_{50} = \frac{\ln(2)}{k}. \quad (1.2)$$

El flujo total transportado J_s es la suma del transporte advectivo y dispersivo:

$$J_s = v\theta C - D\theta \frac{\partial C}{\partial z}, \quad (1.3)$$

Donde v es la velocidad del agua en el poro, θ es el contenido de agua, y C es la concentración residente en la fase líquida. El coeficiente de dispersión D , incluidas la difusión y la dispersión hidrodinámica, es,

$$D = D_0\kappa + \alpha v, \quad (1.4)$$

Donde D_0 es el coeficiente de difusión en el agua, κ es el factor de la matriz del terreno, y α es la longitud de la dispersión. La concentración total es igual a la suma de la concentración de la fase sólida y la concentración de la fase líquida:

$$X = \theta C + \rho Y, \quad (1.5)$$

Donde es la concentración en la fase sólida, y la densidad de la masa seca. Se asume que el pesticida en la fase sólida y la fase líquida está en el equilibrio de acuerdo con un isoterma lineal:

$$Y = K_{oc} f_{oc} C, \quad (1.6)$$

Donde es el coeficiente de adsorción de referencia de la materia orgánica, y la materia orgánica contenida. Al combinar las ecuaciones se obtiene la ecuación de gobierno:

$$R \frac{\partial C}{\partial t} = D \frac{\partial^2 C}{\partial z^2} - v \frac{\partial C}{\partial z} - RkC, \quad (1.7)$$

con el factor de retardación, , definido como

$$R = 1 + \frac{\rho f_{oc} K_{oc}}{\theta}. \quad (1.8)$$

Al llegar a la ecuación (1.7), se han tenido en cuenta las siguientes suposiciones: (i) el contenido de agua, la velocidad del flujo y el coeficiente de dispersión son constantes; (ii) la advección y la dispersión ocurren sólo en dirección vertical; (iii) el factor de retardación es independiente de la concentración; (iv) las transformaciones en las fases líquida y sólida ocurren en la misma tasa.

B. Condiciones iniciales y de contorno

La capa que contiene inicialmente el pesticida se considera que forma parte de la columna, y es incorporada en las condiciones iniciales. El dominio del modelo está dado desde $z = -l$ a $z = \infty$.

Las condiciones iniciales son [4]:

$$C(z, 0) = C_0, \quad -l < z \leq 0, \quad (1.9)$$

$$C(z, 0) = 0, \quad 0 < z \leq \infty, \quad (1.10)$$

Donde es la concentración inicial y es el espesor de la primera capa añadida a la columna de tierra. Son necesarias dos condiciones de contorno. En la definición de la condición de contorno

superior se asume que el agua agregada en esta parte de la columna es libre de pesticida. Además, se asume que no hay declive de la concentración en la profundidad infinita. Por consiguiente, las condiciones de la frontera son dadas como:

$$vC(-l, t) - D \frac{\partial C}{\partial z}(-l, t) = 0, \quad t \geq 0, \quad (1.11)$$

$$\frac{\partial C}{\partial z}(\infty, t) = 0, \quad t \geq 0. \quad (1.12)$$

C. Solución de la ecuación de gobierno

La ecuación de gobierno (1.7) se puede resolver por medio de la transformada de Laplace, técnica usada para ambos dominios semi-infinitos e infinitos. El método de separación de variables es un método alternativo de solución, el cual fue utilizado para verificar los resultados de las soluciones obtenidas con la técnica de la transformada de Laplace. Las variables dimensionales [8] introducidas son:

$\zeta = \frac{vz}{D}$, $\tau = \frac{v^2 t}{RD}$, $\varepsilon = \frac{kRD}{v^2}$, $\lambda = \frac{lv}{D}$,
y el número de Peclet, $p = \frac{Lv}{D}$, donde es la longitud de la columna.

Seguidamente, se presenta la solución analítica de la ecuación (1.7), que está dada por (ver [8]):

$$C(\zeta, \tau) = \frac{1}{2} C_0 \exp(-\varepsilon\tau) \{P + \exp(\zeta + \lambda) Q\}, \quad (1.13)$$

Con,

$$P = \operatorname{erfc} \left[\frac{\zeta - \tau}{2\sqrt{\tau}} \right] - \operatorname{erfc} \left[\frac{\zeta + \lambda - \tau}{2\sqrt{\tau}} \right], \quad (1.14)$$

y,

$$Q = (1 + \tau + \zeta + \lambda) \operatorname{erfc} \left[\frac{\tau + \zeta + \lambda}{2\sqrt{\tau}} \right] \frac{2\sqrt{\tau}}{\sqrt{\pi}} \exp \left[\frac{-(\tau + \zeta + 2\lambda)^2}{4\tau} \right] - (1 + \tau + \zeta + 2\lambda) \operatorname{erfc} \left[\frac{\tau + \zeta + 2\lambda}{2\sqrt{\tau}} \right] + \frac{2\sqrt{\tau}}{\sqrt{\pi}} \exp \left[\frac{-(\tau + \zeta + 2\lambda)^2}{4\tau} \right] \quad (1.15)$$

III. SOLUCIÓN NUMÉRICA DEL MODELO

El problema que nos planteamos es encontrar una función que represente numéricamente la concentración de una sustancia, solución de la ecuación convección-dispersión dada anteriormente,

$$R \frac{\partial C}{\partial t} = D \frac{\partial^2 C}{\partial z^2} - v \frac{\partial C}{\partial z} - RkC, \quad (1.16)$$

cuyas condiciones iniciales se presentan de la siguiente manera: La primera capa, que contiene inicialmente el pesticida, es considerada parte de la columna, y es incorporada en las condiciones iniciales.

Para el caso analítico, el dominio del modelo está dado desde a , y las condiciones iniciales son:

$$C(z, 0) = C_0, \quad -l < z \leq 0, \quad (1.17)$$

$$C(z, 0) = 0, \quad 0 < z \leq \infty. \quad (1.18)$$

Para el caso numérico, varían dichas condiciones como se muestra a continuación y se aprecia en la Figura 1.

$$C(z, 0) = C_0, \quad 0 \leq z \leq l, \quad (1.19)$$

$$C(z, 0) = 0, \quad l < z \leq L, \quad (1.20)$$

De modo que se tiene el valor inicial en el intervalo 0 en el intervalo l . se toma suficientemente grande para que se pueda considerar un dominio semi-infinito para el problema.

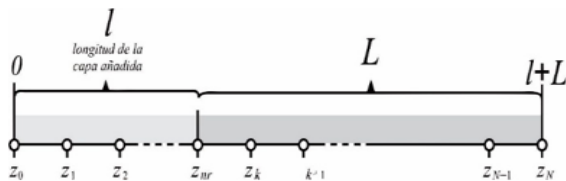


Fig. 1. Representación gráfica del problema numérico

Por consiguiente, las condiciones en los extremos en los distintos instantes de tiempo también deben variar, y se obtienen las ecuaciones siguientes:

$$vC(0, t) - D \frac{\partial C}{\partial z}(0, t) = 0, \quad t \geq 0, \quad (1.21)$$

$$\frac{\partial C}{\partial z}(L, t) = 0, \quad t \geq 0. \quad (1.22)$$

A. Discretización espacial

Para transformar la ecuación en derivadas parciales en un sistema de ecuaciones diferenciales, se procede a discretizar la variable espacial z , en el intervalo $[0, L]$, haciendo subintervalos de longitud h , por lo que la malla se puede definir como $z_0 = 0, z_1 = 0+h, \dots, z_M = L$.

Una vez que se ha discretizado la variable espacial hemos de hacer lo propio con las derivadas espaciales. Por lo que se tiene una aproximación a la primera derivada, utilizando la fórmula de diferencias centradas [9].

$$\left. \frac{\partial C(z, t)}{\partial z} \right|_{z_j} = \frac{C((j+1)\Delta z, t) - C((j-1)\Delta z, t)}{2\Delta z} - \frac{(\Delta z)^2}{6} \left. \frac{\partial^3 C(z, t)}{\partial z^3} \right|_{\zeta_j}, \quad (1.23)$$

Donde $\zeta_j \in [(j-1)\Delta z, (j+1)\Delta z]$.

La aproximación por diferencias centradas anterior tiene un error local de truncamiento de la aproximación de $\frac{\partial C}{\partial z}$, que viene dado por el residuo del desarrollo en serie de Taylor. La aproximación a la segunda derivada por la fórmula de diferencias centradas viene dada por

$$\left. \frac{\partial^2 C(z, t)}{\partial z^2} \right|_{z_j} = \frac{C((j+1)\Delta z, t) - 2C(j\Delta z, t) + C((j-1)\Delta z, t)}{(\Delta z)^2} - \frac{(\Delta z)^2}{12} \left. \frac{\partial^4 C(z, t)}{\partial z^4} \right|_{\zeta_j},$$

Donde $\zeta_j \in [(j-1)\Delta z, (j+1)\Delta z]$.

Una vez hecha esta discretización espacial y definiendo $C_j(t) = C(j\Delta z, t)$, el error local de truncamiento espacial de la aproximación es de $O((\Delta z)^2)$. La ecuación (1.16) discretizada, se escribe

$$R \left(\frac{dC_j}{dt} \right) = \frac{D}{(\Delta z)^2} (C_{j+1} - 2C_j + C_{j-1}) - \frac{v}{2\Delta z} (C_{j+1} - C_{j-1}) - RkC_j - D \left[\frac{(\Delta z)^2}{12} \frac{\partial^4 C(t, \zeta_j)}{\partial z^4} \right] v \left[\frac{(\Delta z)^2}{6} \frac{\partial^3 C(\zeta_j, t)}{\partial z^3} \right],$$

$$j = 1, \dots, L, \quad (1.24)$$

con $\zeta_j \in [z_{j-1}, z_{j+1}]$.

Ahora definamos $c_j(t)$ como la solución del sistema de ecuaciones

$$R \left(\frac{dc_j}{dt} \right) = \frac{D}{(\Delta z)^2} (c_{j+1} - 2c_j + c_{j-1}) - \frac{v}{2\Delta z} (c_{j+1} - c_{j-1}) - Rkc_j, \quad (1.25)$$

Con $j = 1, \dots, L$.

Lo que se espera es que cuando (Δz) sea suficientemente pequeño se cumpla que $c_j(t) \approx C(t, j\Delta z)$.

Luego, con podemos aproximar la ecuación (1.16) de la siguiente manera:

$$\frac{dc_j}{dt} = \frac{D}{R\Delta z^2} (c_{j+1} - 2c_j + c_{j-1}) - \frac{v}{2R\Delta z} (c_{j+1} - c_{j-1}) - kc_j. \quad (1.26)$$

Llamando

$$\alpha = \frac{D}{R\Delta z^2}, \quad \beta = \frac{v}{2R\Delta z},$$

La ecuación (1.26) puede escribirse de forma más simplificada como

$$\frac{dc_j}{dt} = \alpha (c_{j+1} - 2c_j + c_{j-1}) - \beta (c_{j+1} - c_{j-1}) - kc_j = (\alpha + \beta)c_{j-1} - (2\alpha + k)c_j + (\alpha - \beta)c_{j+1}. \quad (1.27)$$

Reemplazando $j=1$ en la ecuación (1.27), se obtiene

$$\frac{dc_1}{dt} = (\alpha + \beta)c_0 - (2\alpha + k)c_1 + (\alpha - \beta)c_2,$$

Análogamente, para valores de $j = 2$ en la ecuación (1.27),

$$\frac{dc_2}{dt} = (\alpha + \beta)c_1 - (2\alpha + k)c_2 + (\alpha - \beta)c_3. \quad (1.28)$$

Evaluando sucesivamente los diferentes valores de j en la ecuación (1.27), se llega al último valor, es decir, $j = L-1$,

$$\frac{dc_{L-1}}{dt} = (\alpha + \beta)c_{L-2} - (2\alpha + k)c_{L-1} + (\alpha - \beta)c_L. \quad (1.29)$$

Se obtiene de esta forma un conjunto de $L-2$ ecuaciones diferenciales ordinarias, que no son suficientes para determinar las incógnitas del problema para la solución aproximada en los nodos de la malla espacial.

Para cerrar el sistema utilizamos las condiciones de contorno, dadas por las ecuaciones (1.21) y (1.22). Al discretizar las derivadas con una fórmula de diferencias centradas en la frontera, van a aparecer valores de la solución en los nodos -1 y $L+1$, c_{-1} y c_{L+1} , dichos nodos deben ser agregados a la malla (denominados nodos ficticios) y evaluar la ecuación diferencial en los nodos 0 y L .

Mediante la aplicación del método de diferencias finitas, que por espacio no se detalla, se obtiene el sistema de ecuaciones diferenciales con incógnitas

$$\frac{dc_0}{dt} = \delta c_0 - 2\alpha c_1$$

$$\frac{dc_j}{dt} = (\alpha + \beta)c_{j-1} - (2\alpha + k)c_j + (\alpha - \beta)c_{j+1},$$

$$j = 1, \dots, L-1$$

$$\frac{dc_L}{dt} = 2\alpha c_{L-1} - (\alpha + k)c_L, \quad (1.30)$$

Si se introduce el vector c , podemos escribir el sistema de ecuaciones (1.30) en forma matricial

$$\vec{c} = \begin{bmatrix} c_0 \\ \vdots \\ c_L \end{bmatrix}, \frac{d\vec{c}}{dt} = \begin{bmatrix} \delta & 2\alpha & 0 & \dots & \dots & \dots & 0 \\ (\alpha + \beta) & -(2\alpha + k) & (\alpha - \beta) & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & (\alpha + \beta) & -(2\alpha + k) & (\alpha - \beta) \\ \dots & \dots & \dots & \dots & 0 & 2\alpha & -(2\alpha + k) \end{bmatrix} \begin{bmatrix} c_0 \\ \vdots \\ c_L \end{bmatrix}$$

Que en forma compacta se representa por

$$\frac{d\vec{c}}{dt} = M\vec{c}.$$

B. Esquema explícito

Para encontrar una discretización completa de esta ecuación necesitamos discretizar también la variable temporal o utilizar un método para resolver el sistema de ecuaciones diferenciales ordinarias que hemos planteado. Comencemos por simplicidad considerando el esquema explícito utilizando el método de Euler [10] y realizando una discretización directa de la derivada temporal por medio de la fórmula de diferencias adelantadas,

$$\frac{dc_j(t)}{dt} = \frac{c_j((n + 1)\Delta t) - c_j(n\Delta t)}{\Delta t} - \frac{\Delta t}{2} \frac{d^2c_j}{dt^2} \Big|_{\zeta_n},$$

Donde . $\zeta_n \in [n\Delta t, (n + 1)\Delta t]$.

Por tanto, la ecuación (1.24) se escribe,

$$R \left(\frac{c_j((n + 1)\Delta t) - c_j(n\Delta t)}{\Delta t} \right) = \frac{D}{(\Delta z)^2} (c_{j+1}(n\Delta t) - 2c_j(n\Delta t) + c_{j-1}(n\Delta t)) - \frac{v}{2\Delta z} (c_{j+1}(n\Delta t) - c_{j-1}(n\Delta t)) - Rkc_j(n\Delta t),$$

Es decir,

$$R \left(\frac{c_j^{n+1} - c_j^n}{\Delta t} \right) = \frac{D}{(\Delta z)^2} (c_{j+1}^n - 2c_j^n + c_{j-1}^n) - \frac{v}{2\Delta z} (c_{j+1}^n - c_{j-1}^n) - Rkc_j^n.$$

Evidentemente, esperamos que se satisfaga y, por lo tanto, al resolver las ecuaciones discretas anteriores estamos construyendo aproximaciones a la solución de la ecuación original en derivadas parciales.

Al introducir $\alpha = \frac{D}{R\Delta z^2}$, $\beta = \frac{v}{2R\Delta z}$, se tiene el sistema en diferencias del método explícito:

$$\begin{aligned} c_0^{n+1} &= c_0^n + \Delta t(\delta c_0^n - 2\alpha c_1^n) \\ c_j^{n+1} &= c_j^n + \Delta t(\alpha + \beta)c_{j-1}^n - \Delta t(2\alpha + k)c_j^n \\ &\quad + \Delta t(\alpha - \beta)c_{j+1}^n \\ j &= 1, \dots, L - 1 \\ c_L^{n+1} &= c_L^n + \Delta t(2\alpha c_{L-1}^n - (2\alpha + k)c_L^n). \end{aligned}$$

En forma matricial, se puede escribir como

$$\vec{c}^{n+1} = \vec{c}^n + \Delta t M \vec{c}^n,$$

A. Esquema implícito

Los esquemas explícitos tienen un coste computacional pequeño en cada paso de tiempo, pero, para ser estables, es necesario trabajar con incrementos de tiempo también pequeños. Un análisis de estabilidad para esquemas explícitos [11], a partir de la teoría de las características para soluciones continuas, lleva a la conclusión que dichos esquemas, para ser estables, deben cumplir la condición de Courant. Dicho lo anterior, debido a que los esquemas explícitos presentan problemas de estabilidad no sólo por el tamaño de paso de discretización, sino por la relación $\frac{\Delta t}{(\Delta z)^2} = r$, donde es llamado el número de Courant, es decir, para algunos valores de Δz y Δt , para evitar este tipo de problemas se pueden utilizar los métodos implícitos. Se tiene el sistema en diferencias del método implícito,

$$\begin{aligned} c_0^n &= c_0^{n+1} - \Delta t(\delta c_0^{n+1} - 2\alpha c_1^{n+1}) \\ c_j^n &= c_j^{n+1} - \Delta t(\alpha + \beta)c_{j-1}^{n+1} + \Delta t(2\alpha + k)c_j^{n+1} \\ &\quad - \Delta t(\alpha - \beta)c_{j+1}^{n+1} \\ j &= 1, \dots, L - 1 \\ c_L^n &= c_L^{n+1} + \Delta t(2\alpha c_{L-1}^{n+1} - (2\alpha + k)c_L^{n+1}). \end{aligned}$$

Lo que se puede evidenciar ahora es que calcular los valores en un punto del espacio y en un instante de tiempo implica utilizar los valores de en otros puntos del espacio en el mismo instante, por lo que se debe resolver en cada paso de tiempo un sistema de ecuaciones que engloba las variables en todos los puntos del espacio en $c_j^{n+1}, j = 0, \dots, L$.

El esquema implícito es de la forma,

$$(I - \Delta t M) \bar{c}^{n+1} = \bar{c}^n,$$

Los esquemas implícitos tienen la ventaja sobre los esquemas explícitos que son incondicionalmente estables independientemente de los pasos y elegidos.

D. Esquema de Crank-Nicholson

Otra posibilidad, si se quiere aproximar la derivada temporal con mayor precisión, es utilizar la regla de medio paso implícita para resolver el sistema de ecuaciones semidiscreto, así obtenemos el denominado esquema de Crank-Nicholson [9], [12] para la ecuación de convección-dispersión

$$\begin{aligned} R \left(\frac{c_j^{n+1} - c_j^n}{\Delta t} \right) &= \frac{D}{2} \left(\frac{c_{j+1}^{n+1} - 2c_j^{n+1} + c_{j-1}^{n+1}}{\Delta z^2} \right) \\ &+ \frac{D}{2} \left(\frac{c_{j+1}^n - 2c_j^n + c_{j-1}^n}{\Delta z^2} \right) \\ &- \frac{v}{2} \left(\frac{c_{j+1}^{n+1} - c_{j-1}^{n+1}}{2\Delta z} + \frac{c_{j+1}^n - c_{j-1}^n}{2\Delta z} \right) \\ &- \frac{Rk}{2} (c_j^{n+1} + c_j^n). \end{aligned}$$

Como la regla de medio paso implícita tiene un error de truncación de orden $O((\Delta t)^2)$, se mejora el comportamiento respecto a los esquemas anteriores que eran de orden $O(\Delta t)$, por lo cual, el error de truncación del esquema de Crank-Nicholson es

$$\tau = O((\Delta t)^2 + (\Delta z)^2).$$

El esquema de Crank-Nicholson es de la forma

$$\left(I - \frac{\Delta t}{2} M \right) \bar{c}^{n+1} = \left(I + \frac{\Delta t}{2} M \right) \bar{c}^n,$$

El esquema de Crank-Nicholson es incondicionalmente estable [12], [13], independientemente de los pasos Δz y Δt elegidos.

IV. RESULTADOS NUMÉRICOS DE LA ECUACIÓN CONVECCIÓN-DISPERSIÓN

En este apartado se presentan los resultados numéricos obtenidos con los esquemas explícito e implícito. Para la simulación numérica que se presentan en este capítulo, se utilizan los valores de los parámetros que se presentan en la Tabla 1 [6]. Que constituyen un problema típico de transporte de pesticida en un columna de suelo. Esto permitirá evaluar el funcionamiento de los métodos propuestos en este tipo de problemas.

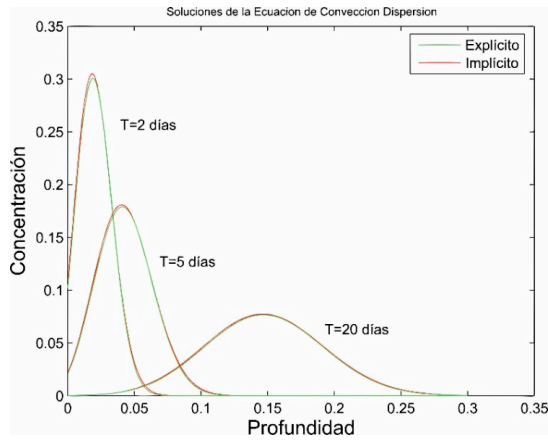
Tabla 1. Valores de los parámetros para las simulaciones [1]

Parámetro	Valor	Unidades
l	variable	m
K_{oc}	0.100	$m^3 kg^{-1}$
T_{50}	100	d
θ	0.40	-
$v\theta$	0.0125	$m d^{-1}$
D_0	4×10^{-5}	$m^2 d^{-1}$
α	0.005	m
f_{oc}	0.01	-
ρ	1400	$kg m^{-3}$
κ	0.34	-

Fuente: Elaboración propia.

La Figura 2 muestra la simulación del modelo bajo los esquemas explícito e implícito, con $\Delta t = 0.05$ (días), tiempo final T para 2, 5 y 20 días, $\Delta z = 0.0001 m$ y el espesor de la capa añadida $l = 0.01 m$. Como se puede evidenciar, los resultados varían levemente, sin embargo, el esquema implícito resulta ser estable para todo valor de los parámetros de discretización.

Fig. 2. Simulación de la ecuación convección-dispersión: esquema implícito vs explícito



Para las gráficas que aparecen en esta sección (Figura 3 y Figura 4), el eje de las x representa la profundidad en m (metros) y el eje de las y representa la concentración en $mol\ m^{-3}$.

Como se puede evidenciar, el esquema implícito es un esquema estable para todo valor de los parámetros de discretización. Sin embargo, el costo computacional es más alto que el del esquema explícito, debido a la utilización de memoria (por la necesidad de almacenar la matriz) por el número de operaciones que hay que realizar, al tener que resolver un sistema de ecuaciones en cada iteración de tiempo.

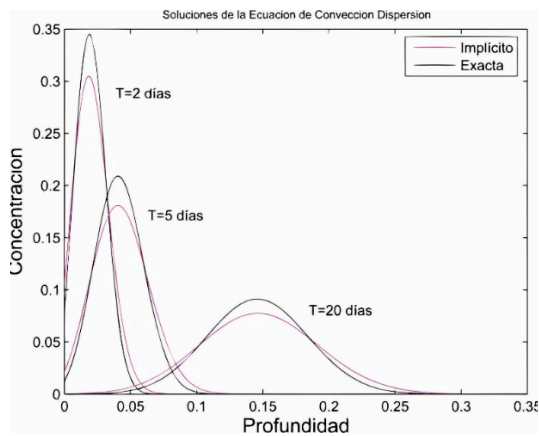


Fig. 3. Simulación de la ecuación convección-dispersión con el esquema implícito con $\Delta t = 0.05$ (días), tiempo final T 2, 5 y 20 días, $\Delta z = 0.0001\ m$ y el espesor de la capa añadida $l = 0.01\ m$. Solución exacta en el tiempo final T 2, 5 y 20 días y valor de la capa añadida $l = 0.01\ m$.

Sin embargo, y de acuerdo con los resultados obtenidos, se puede notar que de los tres esquemas desarrollados el que mejor aproxima los valores obtenidos con la solución exacta es el esquema de Crank-Nicholson, se puede ver en los gráficos que este esquema hace una buena aproximación de la simulación exacta, incluso en los picos, donde es más notorio el error de aproximación.

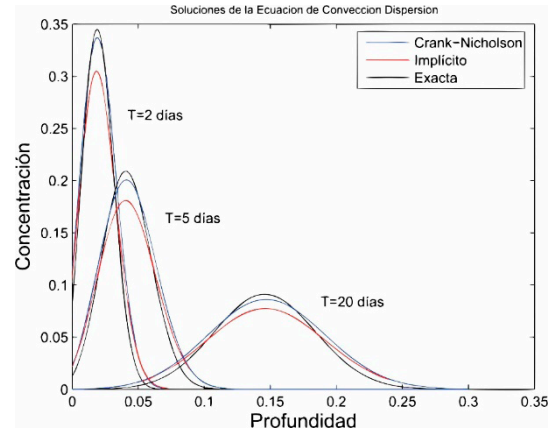


Fig. 4. Simulación de la ecuación convección-dispersión con el esquema implícito con $\Delta t=0.05$ (días), $\Delta t=0.0001$ (m), tiempo final T 2, 5 y 20 días, el espesor de la capa añadida $l=0.01\ m$. Simulación de la ecuación convección-dispersión con el esquema Crank-Nicholson con $\Delta t=0.05$ (días), $\Delta t=0.0001$ (m), tiempo final T 2, 5 y 20 días, el espesor de la capa añadida $l=0.01\ m$. Solución exacta en el tiempo final T 2, 5 y 20 días y valor de la capa añadida $l=0.01\ m$.

V. CONCLUSIONES

Se soluciona la ecuación de advección-convección en forma numérica utilizando los métodos explícito e implícito sobre el problema propuesto por Freijer [4], con las condiciones de contorno e iniciales allí propuestas. Se comparan los resultados obtenidos numéricamente con la solución analítica y se encuentra que los métodos propuestos son buenas aproximaciones de la solución analítica. El método implícito presenta mejor estabilidad para cualquier valor en los parámetros de discretización, pero a un mayor costo computacional. En los

trabajos futuros se espera probar la eficiencia de estos métodos resueltos utilizando computación de alto rendimiento, con lo cual el costo de cómputo ya no sería un problema.

REFERENCIAS

- [1] S. Dardouri and J. SGhaier, “Adsorption characteristics of layered soil as delay barrier of some organic contaminants: Experimental and numerical modeling”, *Environ. Modell. Software*, vol. 110, pp. 95-106, 2018.
- [2] P. Tongyan and T. Miao, “Contamination of roadside soils by runoff pollutants: A numerical study”, *Transp. Geotech.*, vol. 2, pp. 1-9, 2015.
- [3] D. Archundia, C. Duwig, L. Spadini, M. C. Morel, B. Prado, M. P. Perez, V. Orsag and J. M. F. Martins, “Assessment of the Sulfamethoxazole mobility in natural soils and of the risk of contamination of water resources at the catchment scale”, *Environ. Int.*, vol. 130, September 2019.
- [4] J. L. Freijer, E. J. M. Veling, S. M. Hassanizadeh, “Analytical solutions of the convection-dispersion equation applied to transport of pesticides in soil columns”, *Environ. Modell. Software*, vol. 13, pp. 139-149, 1998. Available: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.495.1059&rep=rep1&type=pdf>
- [5] BBA, “Seepage behaviour of plant protection products”, *Guidelines for the official testing of plant protection products*. Part IV 4-2, Braunschweig, Germany, 1986.
- [6] J. I. Freijer, S. Q. Broerse, S. M. Hassanizadeh, A. M. A. Van der Linden, E. J. M. Veling, “Column leaching experiments for aged residues of pesticides: Interpretation and criteria”, Report no. 715801004, Bilthoven, The Netherlands: National Institute of Public Health and the Environment, 1995.
- [7] United States Environmental Protection Agency (EPA), Proposed guidelines for registering pesticides in the United States, Chapter (b) of Section 163, pp. 62-69, Federal Register 43(132), 29719, 1978.
- [8] M. Th. Van Genuchten and W. J. Alves, *Analytical solutions of the one-dimensional convective-dispersive solute transport equation*. USDA Technical Bulletin No. 1661, Washington DC: US Government Printing Office, 1982.
- [9] R. Haberman, *Ecuaciones en derivadas parciales con series de Fourier y problemas de contorno*. 3.^a ed., Madrid: Pearson Educación, 2003.
- [10] W. F. Ames, *Numerical Methods for Partial Differential Equations*. 2nd ed., Computer Science and Applied Mathematics, Orlando, Florida: Academic Press, 1970.
- [11] L. F. León and A. P. Martínez, “Stability Criterion for Explicit Schemes (Finite-Difference Method), on the Solution of the Advection-Diffusion Equation”, in *Proc. of the VIII Int. Conf. on Comput. Methods on Water Resources*, Venice, Italy, 1990, pp. 381-386.
- [12] R. Burden and J. Douglas Faires, *Numerical Analysis*. Boston, USA: CENGAGE Learning, 2015.
- [13] C. D. S. Tomlin, *The pesticide manual*. Farnham, Surrey, UK: British Crop Protection Council, 2000.

