



CoNCISa 2024

Sociedad Venezolana de Computación

Memorias de la Décima Conferencia Nacional
de Computación, Informática y Sistemas

18 al 20 de noviembre de 2024
Universidad Central de Venezuela
Caracas, Venezuela





Memorias de la Décima Conferencia Nacional de Computación, Informática y Sistemas

CoNCISA 2024

18 al 20 de noviembre de 2024
Universidad Central de Venezuela
Caracas - Venezuela

Editores

Eric Gamess, Jesús Pérez



Depósito Legal: DC2024001071
ISBN: 978-980-7683-08-1



@ 2024 - Sociedad Venezolana de Computación
<https://svc.net.ve>

Memorias de la Décima Conferencia de Computación, Informática y Sistemas (CoNCISa 2024).

Depósito Legal: DC2024001071
ISBN: 978-980-7683-08-1

Las contribuciones de CoNCISa 2024 han sido publicadas en la página web de la conferencia.
<https://concisa.net.ve/2024>

El diseño gráfico de la portada y contraportada fue realizado por Frank Rodríguez.

Los conceptos y puntos de vista expresados en los trabajos publicados en estas memorias representan las opiniones personales de los autores y no reflejan el juicio de los editores, de las universidades participantes, o de la Sociedad Venezolana de Computación.

Al someter sus contribuciones a CoNCISa, los autores de los trabajos publicados en estas memorias certifican que todo el material sujeto a copyright usado en sus artículos, incluyendo pero no limitado a figuras, tablas, extractos textuales y muestras de datos ha sido apropiadamente autorizado a través de permisos escritos, que son debidamente documentados y disponibles para verificación.

Pagad a todos lo que debéis:
al que tributo, tributo; al que impuesto, impuesto; al que respeto, respeto; al que honra, honra.
No debáis a nadie nada, sino el amaros unos a otros; porque el que ama al prójimo, ha cumplido la ley.
(Romanos 13:7-8)

Prólogo

La Sociedad Venezolana de Computación (SVC), con su firme propósito de promover la formación y el desarrollo científico-tecnológico en el área de informática en nuestro país, organiza eventos científicos y educativos de alto nivel como lo son la Escuela Venezolana de Informática (EVI) y la Conferencia Nacional de Computación, Informática y Sistemas (CoNCISa).

EVI es un evento en el que expertos nacionales e internacionales dictan cursos especializados y pertenecientes a diferentes dominios de las ciencias computacionales. Estos cursos cubren temas de interés para la disciplina Informática y son complementarios a las materias y temas que figuran en los currículos de las carreras de computación, informática y sistemas ofrecidas en el país.

CoNCISa es una conferencia amplia que abarca áreas relacionadas con las ciencias de computación, informática y sistemas, cuyo objetivo principal es promover intercambios periódicos sobre experiencias académicas, tecnológicas y de investigación, buscando impulsar el desarrollo del área y crear lazos estrechos de cooperación a nivel nacional. En este sentido, CoNCISa provee espacios para la presentación y discusión de trabajos, resultados y tecnologías, organizados en sesiones de investigación y foros especializados. Además, la conferencia complementa su importancia a nivel nacional, al invitar renombrados investigadores en las áreas de computación, informática y sistemas para exponer sobre los avances, tendencias y direcciones de investigación con mayor auge a nivel internacional.

Este año, con el apoyo de las instituciones académicas Universidad Central de Venezuela (UCV), Universidad Simón Bolívar (USB), Universidad Católica Andrés Bello (UCAB) y Universidad de Los Andes (ULA), se ha organizado EVI y CoNCISa 2024, en su décima edición, del 18 al 20 de noviembre de 2024.

CoNCISa 2024 recibió dos tipos de contribuciones: (1) artículos cortos y (2) artículos largos. Después de una selección rigurosa de todos los trabajos recibidos, se aceptaron tres artículos cortos y cuatro artículos largos para presentación oral, que están incluidos en esta memoria. Todos los trabajos enviados fueron evaluados por al menos tres miembros del Comité de Programa, aplicando adecuadamente los criterios de valoración establecidos.

La realización exitosa de esta conferencia corresponde al esfuerzo de reconocidos investigadores, profesores y estudiantes del área a nivel nacional e internacional, quienes contribuyeron con su tiempo y dedicación en las actividades de organización, evaluación y edición. Así, nuestro reconocimiento va a todos los autores que confiaron en CoNCISa como una conferencia merecedora de sus excelentes contribuciones, producto de sus trabajos de investigación; a los miembros de los comités de programa y organizador, conferencistas, expositores de tutoriales; a los profesores y estudiantes, y en general, a todos aquellos que aportaron su valioso apoyo para llevar a buen término esta importante responsabilidad, manteniendo en alto la relevancia de nuestra conferencia nacional; a todos, muchas gracias.

Yosly Hernández
Universidad Central de Venezuela

Jesús Pérez
Universidad de Los Andes

Presidentes del Comité de Programa de CoNCISa 2024

La SVC (Sociedad Venezolana de Computación) está comprometida con el impulso de una nueva generación académica y profesional en nuestra área de saber para el desarrollo del país

Comité de Programa

Presidentes del Comité de Programa

Yosly Hernández	Universidad Central de Venezuela – Venezuela
Jesús Pérez	Universidad de Los Andes – Venezuela

Miembros del Comité de Programa

Junior Altamiranda	Universidad de Los Andes – Venezuela
Alejandro Amaro	Universidad Nacional Experimental de la Gran Caracas – Venezuela
Yudith Cardinale	Universidad Simón Bolívar – Venezuela
César Collazos	Universidad del Cauca – Colombia
Eric Gamess	Jacksonville State University – USA
Luis Hernández	Universidad Central de Venezuela – Venezuela
Vanessa Leguízamo	Universidad Central de Venezuela – Venezuela
Jonás Montilva	Universidad de Los Andes – Venezuela
Dinarle Ortega	Universidad Católica Andrés Bello – Venezuela
Jaime Parada	Universidad Central de Venezuela – Venezuela
Robinson Rivas	Universidad Central de Venezuela – Venezuela
Franklin Sandoval	Universidad Central de Venezuela – Venezuela
Andrés Sanoja	Universidad Central de Venezuela – Venezuela
Eugenio Scalise	Universidad Central de Venezuela – Venezuela
Leonid Tineo	Universidad Simón Bolívar – Venezuela

Comité Organizador

Presidentes del Comité Organizador

Yosly Hernández	Universidad Central de Venezuela – Venezuela
Jesús Pérez	Universidad de Los Andes – Venezuela
Andrés Sanoja	Universidad Central de Venezuela – Venezuela

Miembros del Comité Organizador

Concettina Di Vasta	Universidad Central de Venezuela – Venezuela
Eric Gamess	Jacksonville State University – USA
Jaime Parada	Universidad Central de Venezuela – Venezuela
Paul Quijada	Universidad Central de Venezuela – Venezuela
Keyla Rivas	Universidad Central de Venezuela – Venezuela
Rosseline Rodríguez	Universidad Simón Bolívar – Venezuela
Franklin Sandoval	Universidad Central de Venezuela – Venezuela
Leonid Tineo	Universidad Simón Bolívar – Venezuela

Tabla de Contenido

Prólogo	iii
Comité de Programa	iv
Comité Organizador	v
Tabla de Contenido	vi
1. Sesión de Artículos Cortos	
1.1. Modelo para la Construcción de Políticas Públicas con el Uso de Tecnologías Emergentes bajo el Enfoque 4.5	2-5
Ángel Gil, Marcelita Arroyo, Jesús García	
1.2. Pautas Lógicas para el Diseño Computacional de Sistemas Críticos Basadas en el Paradigma de Fitch y Muckin	6-9
Miguel Torrealba	
1.3. Proceso de Análisis del Dominio para Arquitecturas de Referencias IMFS & VCD Modelado con SPEM	10-15
Aleidys Arraiz, Yusneyi Carballo	
2. Sesión de Artículos Largos	
2.1. Customizing Software Development Methods: A Process Model Approach	17-24
Judith Barrios, Jonás Montilva	
2.2. Estrategias para Impulsar el Ecosistema Emprendedor Tecnológico Venezolano: Análisis DOFA Ampliado	25-31
José Fernández, Yusneyi Carballo	
2.3. Importancia de la Catalogación Digital en la Conservación del Patrimonio Natural: Ficoflora Venezuela, Un Caso de Estudio	32-43
Yusneyi Carballo, Santiago Gómez	
2.4. Análisis de Sentimientos Migratorios en Publicaciones de Venezolanos en Medios Sociales	44-53
Livia Borjas, Mauricio Morales, Miguel Zamora	
Índice de Autores	55



Sesión de Artículos Cortos

Modelo para la Construcción de Políticas Públicas con el Uso de Tecnologías Emergentes bajo el Enfoque 4.5

Ángel Gil¹, Marcelita Arroyo², Jesús M. García³
aegp17@gmail.com, marcearroyo_2@yahoo.com, jmgarcia@unet.edu.ve

¹ Laboratorio de Prototipos, Universidad Nacional Experimental del Táchira (UNET), TEPUY I+D, Quito, Ecuador

² Red GESODEL, Quito, Ecuador

³ Laboratorio de Prototipos, Universidad Nacional Experimental del Táchira (UNET), San Cristóbal, Venezuela

Resumen: Como herramienta de transformación de la calidad de vida de la ciudadanía, las políticas públicas establecen los objetivos para incidir de forma positiva en la sociedad y contribuir con la solución de los problemas o situaciones que emergen de las dinámicas sociales de los países, esto mediante el uso de los medios del estado. La capacidad de co-crear estas políticas en total simbiosis con la ciudadanía es un reto constante para los gobiernos, y en la actualidad el uso de las nuevas tecnologías se presenta como un medio con alto valor para la co-creación de políticas efectivas y adaptadas a la realidad de los espacios que buscan afectar, específicamente desarrollando políticas fundamentadas en datos, con apoyo de tecnologías emergentes para su diseño, implementación y evaluación constante. El enfoque 4.0, muy orientado a la industria y a la modernización de los procesos, puede ser aplicado en estos espacios de forma tal, que permita modernizar y, sobre todo, democratizar la construcción de las políticas que rigen los espacios de participación en pro de mejorar la calidad vida de la ciudadanía, pasando de herramientas reactivas a herramientas proactivas. Por ello, se propone un enfoque 4.5 para la construcción de políticas públicas desde un enfoque colaborativo e incluyente.

Palabras clave: Enfoque 4.5; Política Pública; Inteligencia Artificial; Tecnologías Emergentes.

1. INTRODUCCIÓN

Tradicionalmente el enfoque para la construcción de las políticas públicas (PP) ha sido vertical, una construcción de arriba hacia abajo, donde la verdad del estado es la que predomina [1]; esto ha venido cambiando de forma tal, que ahora involucra a diversos actores sociales, quienes finalmente son los afectados por las decisiones que se toman. Este enfoque horizontal o de co-creación, ha permitido la generación de espacios de discusión que conllevan a la construcción de políticas participativas. Ahora bien, estos espacios son complejos, por ello requieren herramientas que les permitan mejorar su efectividad y finalmente, contribuir con soluciones y alternativas reales a los problemas de la ciudadanía.

El enfoque 4.0, donde la tecnología es utilizada para automatizar las distintas etapas de un proceso industrial, es ampliado en esta investigación para permitir que, a partir de la interacción entre la población y en un enfoque bottom-up, se construyan entornos de inclusión y democratización fundamentados en herramientas inteligentes, que faciliten la interacción entre ciudadanía y el estado, permitiendo el surgimiento de políticas que estén adaptadas a la realidad de los territorios, así como un mejor aprovechamiento de los recursos asignados y la generación de alternativas de solución que perduren en el tiempo y evolucionen conforme a las propias dinámicas y modos de vida locales. Se propone el enfoque 4.5 desde una perspectiva técnico - social centrado en la búsqueda de una mejor calidad de vida para la ciudadanía a través de la integración de las políticas con herramientas emergentes, integradas y consistentes.

Específicamente, en este trabajo se presenta un modelo para la construcción de políticas públicas 4.5, fundamentado en el ciclo de la política pública [2]: diagnóstico, diseño, implementación y evaluación, en conjunto con tres dimensiones: interpretación de la realidad, transformación de la realidad y evaluación de la realidad transformada.

2. ESTADO DEL ARTE

En esta sección se presentan investigaciones relacionadas con el uso de la Inteligencia Artificial (IA) en el gobierno y en el marco de las PP. En [3] se describe cómo en cada fase del ciclo de las PP, las herramientas emergentes pueden dar soporte para mejorar los procesos, se mencionan además las posibles aplicaciones entre las que se encuentra el *gobierno abierto* y la *administración pública*. Por otra parte, se hace énfasis en las implicaciones negativas del uso de la tecnología en un proceso ciudadano, como son los errores que puedan aparecer en las distintas etapas y que puedan afectar el resultado, además, se describe el riesgo de los sesgos en los algoritmos.

En [4] se presenta el problema de incertidumbre en el uso de las tecnologías; en este caso específico, el uso de la IA que puede generar una incertidumbre que se traspasa a los responsables de tomar decisiones sobre las PP. Por otra parte, en la medida que las tecnologías han ido apareciendo, han sido objeto de regulaciones por parte de los gobiernos que buscan gestionar su uso correcto; en [5] se comenta sobre cómo estas regulaciones pudieran afectar las soluciones en espacios asociados a las PP y sobre la necesidad de no socavar la innovación.

En este sentido, actualmente las IA están siendo sometidas a regulaciones como parte de las PP, por tanto, en [6] se discute sobre cómo las leyes relacionadas con la IA conllevan a un

control de riesgo efectivo, y generan implicaciones junto con responsabilidades individuales y colectivas en el mundo actual, donde la interconexión es el centro de los procesos. Asimismo, en [7] se discute cómo las PP deberían afrontar en el futuro, la aparición de una IA “super inteligente” junto con los nuevos desafíos asociados, que deberán abordarse desde las normativas de los estados.

En este contexto, en [8] se analiza la necesidad recurrente y materializada en los años recientes, de regular tecnologías como la IA en función de los problemas futuros que pueda causar, y no fundamentado en las prestaciones actuales de las herramientas disponibles. Si este asunto se relaciona con las PP, quedan muchos temas por discutir y adaptar a los tiempos actuales.

Finalmente, en [9] se discute sobre las implicaciones de la implementación de IA en los procesos de gobierno, donde parece estar claro que la tecnología mejora la interacción con los ciudadanos, la prestación de servicios y la construcción de PP, pero esto amerita una transformación tecnológica a todos los niveles y la interdisciplinariedad en los profesionales de los diferentes sectores de gobierno.

3. MODELO 4.5 PARA LA CONSTRUCCIÓN DE POLÍTICAS PÚBLICAS

El modelo 4.0 en la industria busca combinar distintas tecnologías de producción de bienes y servicios, con herramientas inteligentes, integradas con las personas, los activos y la organización, con el objetivo de mejorar los procesos productivos [10]. En este trabajo se propone un modelo 4.5 que se orienta hacia la integración de herramientas inteligentes en la construcción de PP, las cuales representan un aspecto fundamental para mejorar la convivencia y calidad de vida de las personas en su entorno social; por tanto, se busca mejorar el proceso social y adecuarlo a las necesidades y realidades de las sociedades latinoamericanas.

En este sentido, se presenta en este trabajo un modelo para la construcción de PP, tomando como base el ciclo de desarrollo de las PP [2]:

- **Diagnóstico:** con el objetivo de identificar las oportunidades o problemas presentes en un entorno social.
- **Diseño:** elaboración de la PP a través de la coordinación con distintos actores, enfocados en la generación de una solución a la necesidad identificada.
- **Implementación:** fase donde se ejecuta la PP a través de las herramientas dispuestas para ello por el estado.
- **Evaluación:** es necesario evaluar el impacto de la PP en función de los resultados esperados y en búsqueda de la mejora continua.

Al abordar este ciclo, existen distintos desafíos [11] en la construcción de PP, relacionados con aspectos críticos que se deben enfrentar para alcanzar un proceso de construcción efectivo:

- El acceso a la información de forma oportuna puede verse restringido por las condiciones propias de los territorios.
- La veracidad de la información (ruido) afecta el análisis y el diagnóstico.

- La información puede estar depreciada en el tiempo o con sesgo.
- Los actores involucrados pueden presentar debilidades en los temas base para la construcción efectiva de PP.
- Los mecanismos de participación ciudadana pueden verse afectados por las condiciones del territorio o por las condiciones propias de cada grupo social.
- La ejecución de los presupuestos asignados, el seguimiento adecuado, el control, impactan en la eficiencia de la PP.

Ahora, sabiendo que el fin de la PP es transformar la realidad de un grupo social para bien, mejorando de la calidad de vida de las personas y solucionando los problemas como principio fundamental; se propone en este modelo, la implementación de tres dimensiones desde el punto de vista técnico-social, que permitan coadyuvar en el desarrollo de PP al permear en cada una de las fases su ciclo:

- **Interpretación de la realidad:** diagnóstico efectivo que debe tomar en cuenta la realidad presente en los territorios a impactar y la realidad de las comunidades.
- **Transformación de la realidad:** una PP efectiva debe ser capaz de transformar la realidad de forma positiva.
- **Evaluación de la realidad transformada:** la evaluación permite identificar el impacto y la sostenibilidad en el tiempo de la PP así como la posibilidad de mejora de esa realidad transformada.

Este modelo busca integrar el uso de IA, analítica de datos y otras herramientas emergentes en cada uno de los procesos de la construcción de las PP considerando el cambio de la realidad en los actores, para extender bajo un enfoque 4.5 el ciclo de construcción de la PP, como se muestra en la Figura 1.

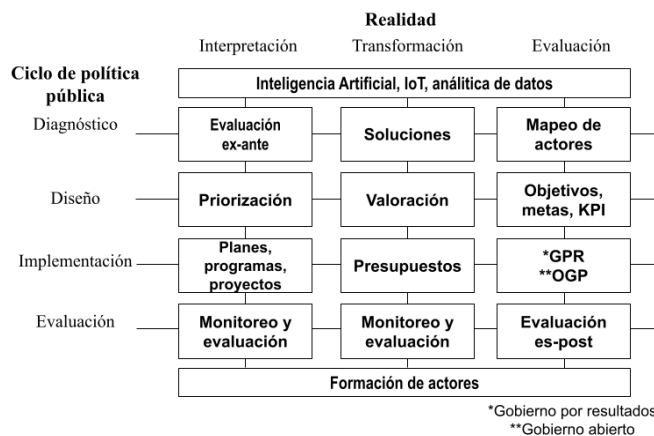


Figura 1: Modelo 4.5 para la Construcción de Políticas Públicas

Por otra parte, el modelo propuesto se soporta en dos ejes fundamentales:

1. **Uso de herramientas tecnológicas emergentes:** En la actualidad el uso de la IA, internet de las cosas (IoT), y analítica de datos es cada vez más común en distintos entornos de la sociedad. En este sentido, la adopción pública de los modelos de lenguaje de gran tamaño (LLM) ha representado un salto importante para la difusión y uso

de la IA por gran parte de la ciudadanía, por tanto, su implementación en diversos espacios ya es una necesidad: cada día se generan gran cantidad de datos que deben ser recopilados, preprocesados y analizados, para obtener información de calidad a partir de los mismos. En el caso concreto de las PP, la información es fundamental en todo su ciclo, es por ello que el modelo propuesto toma como base estas herramientas.

2. Formación de los actores que intervienen en la construcción de PP: La adopción de herramientas tecnológicas trae consigo un aspecto clave: la formación de los actores que intervienen en los procesos donde son utilizadas, esto con el fin de que la ciudadanía esté no solo involucrada en el proceso como fuente de información o como beneficiaria del proceso ejecutado, sino que también sea capaz de conocer y manejar distintos conceptos, que permitan generar un entorno transparente y de confianza. Así los ciudadanos serán partícipes en todas las fases del ciclo de construcción de las PP, teniendo como principio la colaboración, la cual es un pilar del modelo de gobierno abierto.

Sobre estos dos ejes, se construye el modelo y cada una de sus fases, donde intervienen procesos de diagnóstico, identificación, priorización, planificación, construcción de indicadores, evaluación, entre otros; que en conjunto, permitirán construir un entorno adecuado para la co-creación, implementación y seguimiento de PP que den solución a los problemas de las comunidades beneficiarias de las mismas.

4. HERRAMIENTAS EMERGENTES EN PP

La construcción de políticas públicas sigue un ciclo clásico de diseño e implementación: diagnóstico, planificación, ejecución y evaluación [12]; cada una de estas etapas son ejecutadas según las necesidades y recursos disponibles por los actores involucrados, y de acuerdo con las normas establecidas por las instituciones del estado. En este ciclo se pueden identificar de forma general algunos problemas:

- Recopilación de datos: la toma de datos puede presentar un problema debido a la heterogeneidad de los territorios y comunidades, donde se pueden presentar deficiencias como accesibilidad, conexión a internet o telefonía, problemas sociales, vulnerabilidad, entre otros.
- Confiabilidad de los datos: lo mencionado anteriormente hace necesario el preprocesamiento de los datos y contar con herramientas que permitan obtener datos confiables y con el menor sesgo posible.
- Generación de información: el análisis de esos datos y la información generada es fundamental, además de la interacción con los distintos actores para que la información sea coherente con la realidad.
- Consolidación de información: consolidar la información y construir fuentes de datos únicas, es de suma importancia. Dichas fuentes deben ser accesibles por los diferentes actores, además es preciso que sean confiables y seguras, esto es un aspecto clave en la construcción de buenas PP.
- Seguimiento: la capacidad de llevar a cabo procesos de seguimiento dentro de los territorios y en conjunto con la

ciudadanía permite que las PP sean ejecutadas de forma efectiva.

- Evaluación: la evaluación constante, la detección de desviaciones o de nuevas necesidades, permiten actuar de forma oportuna para aplicar los correctivos necesarios en pro del beneficio de la ciudadanía.

Vistos estos problemas, se propone el uso de distintas herramientas que permitan abordar de una forma integral el ciclo de construcción de PP y adoptar el modelo 4.5:

1. IoT en la toma de datos en los territorios.
2. Construcción de herramientas de adquisición de datos, que permitan de forma inteligente realizar un preprocesamiento previo de los datos recolectados.
3. Uso de Data lake para almacenamiento de datos.
4. Uso de algoritmos inteligentes para el procesamiento de los datos recolectados.
5. Uso de Inteligencia de negocios (BI) para el análisis de los datos recolectados.
6. Uso de sistemas expertos para la construcción de propuestas de PP según la información obtenida.
7. Uso de herramientas cooperativas para la disponibilidad de los datos.
8. Construcción de herramientas de seguimiento y control.
9. Implementación de IoT para el seguimiento y control de acuerdo a las necesidades de los territorios.
10. Análisis en tiempo real de los datos recolectados y la identificación de patrones de alerta.
11. Implementación de herramientas que faciliten la transparencia de la información, así como procesos de auditoría adecuados.

5. INTEGRACIÓN DEL MODELO EN UN EJEMPLO REAL

Un caso de aplicación relacionado con el diseño de una PP podría centrarse en el uso IA y análisis de datos para mejorar la gestión de la movilidad urbana. Este enfoque permite desarrollar políticas que respondan de manera más precisa y dinámica a las necesidades de los ciudadanos en tiempo real.

5.1 Contexto

Las ciudades grandes enfrentan un crecimiento constante de la población urbana, lo que trae consigo desafíos complejos de movilidad, tales como congestión, contaminación y accidentes. Tradicionalmente, los enfoques de planificación urbana se fundamentan en análisis de datos estáticos y modelos históricos que pueden resultar insuficientes para responder a la naturaleza cambiante de las ciudades.

5.2 Aplicación de la Tecnología Emergente

En este caso, el gobierno implementa una política de movilidad urbana basada en IA y análisis de Big data para monitorear y gestionar el flujo de tráfico en tiempo real. Este tipo de tecnología puede recolectar y analizar datos de distintas fuentes, como sensores de tráfico, GPS en vehículos, cámaras de vigilancia y dispositivos móviles, que permiten crear modelos predictivos de tráfico.

5.3 Ejemplo de Impacto en Política Pública

1. *Política de Transporte Público Inteligente:* con la información obtenida, la PP podría optimizar las rutas y frecuencias del transporte público en tiempo real. Si un sensor detecta un aumento repentino en el tráfico en una zona específica, el sistema podría redirigir autobuses o trenes para mejorar la cobertura y minimizar la congestión.
2. *Mejora de la Seguridad Vial:* mediante el análisis de datos históricos y en tiempo real, la IA puede identificar áreas de alta accidentalidad y sugerir intervenciones para mejorar la infraestructura o las políticas de señalización en puntos específicos. De esta manera, la PP puede reducir la siniestralidad al priorizar las áreas más críticas.
3. *Sostenibilidad y Reducción de Emisiones:* los datos pueden ayudar a establecer restricciones dinámicas de circulación en función de la calidad del aire. Por ejemplo, si los sensores detectan un incremento en los niveles de contaminación en una zona determinada, la PP puede activar alertas o restricciones temporales de tráfico para ciertos tipos de vehículos.

5.4 Resultados Esperados

- Eficiencia en la distribución del transporte urbano.
- Reducción de la contaminación y mejora de la calidad del aire.
- Disminución de accidentes viales en zonas críticas.
- Mayor satisfacción ciudadana al responder a sus necesidades de movilidad en tiempo real.

En este ejemplo de PP utilizando IA y Big data se destaca cómo las tecnologías emergentes de tipo 4.5 pueden transformar la gestión urbana para adaptarse a las necesidades cambiantes y mejorar la calidad de vida en las ciudades.

5.5 Resumen

La IA y Big data constituyen tecnologías que pueden contribuir con un mejoramiento en la movilidad de las grandes ciudades, mitigando las problemáticas de congestión, contaminación y accidentes. El impacto del uso de estos mecanismos se vería reflejado en contar con políticas de transporte público inteligente, mejora de la seguridad vial y reducción de emisiones, lo que contribuye a una movilidad sostenible. Estas tecnologías permitirían gestionar el flujo del tráfico en tiempo real utilizando sensores de tráfico, GPS en vehículos, cámaras de vigilancia y dispositivos móviles, que permiten crear modelos predictivos de tráfico.

6. CONSIDERACIONES FINALES

El modelo propuesto para la co-creación de PP 4.5 resalta la importancia de integrar tecnologías emergentes, como la IA,

IoT y análisis de datos avanzado, en todas las fases del ciclo de las PP. Esto no solo influye en la eficiencia, efectividad y sostenibilidad de las políticas, sino también asegura su relevancia y adaptabilidad a las realidades locales y modos de vida. Esta integración tecnológica facilita y promueve la participación de la ciudadanía y la toma de decisiones basadas en datos, transformando las PP desde herramientas reactivas hasta proactivas, las cuales son capaces de anticipar y responder a las necesidades cambiantes de la sociedad y contribuyendo al fortalecimiento de la confianza mediante el trabajo colaborativo.

REFERENCIAS

- [1] A.-N. Roth, «Las Políticas Públicas y la Gestión Pública: Un Análisis desde la Teoría y la Práctica,» *Estudios de la Gestión: revista internacional de administración*, nº 5, pp. 223–229, 2019.
- [2] S. N. d. Planificación, «Guía Metodológica para la Formulación de Política Pública,» Febrero 2022. <https://oei.int/GuiaMetodologicaParaLaFormulacionDeLaPoliticaPublica-17-02-2022.pdf>.
- [3] D. Valle-Cruz, E. Ruvalcaba-Gomez, R. Sandoval-Almazan, and J. Criado, «A Review of Artificial Intelligence in Government and its Potential from a Public Policy Perspective,» de *20th Annual International Conference on Digital Government Research*, Dubai, 2019.
- [4] M. Nordström, «AI Under Great Uncertainty: Implications and Decision Strategies for Public Policy,» *AI & Society*, vol. 37, pp. 1703–1714, 2022.
- [5] A. Thierer, A. Castillo O'Sullivan, and R. Russell, «Artificial Intelligence and Public Policy,» 22 Agosto 2017. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3021135.
- [6] H. Zech, «Liability for AI: Public Policy Considerations,» *ERA Forum*, vol. 22, pp. 147–158, 2021.
- [7] N. Bostrom, A. Dafoe, and C. Flynn, «Public Policy and Superintelligent AI: A Vector Field Approach,» de *Ethics of Artificial Intelligence*, S. M. Liao, Ed., New York, Oxford University Press, 2020.
- [8] P. M. Krafft, M. Young, M. Katell, K. Huang, and G. Bugingo, «Defining AI in Policy versus Practice,» de *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, New York, 2020.
- [9] D. Valle-Cruz, J. Ignacio Criado, R. Sandoval-Almazán y E. A. Ruvalcaba-Gomez, «Assessing the Public Policy-cycle Framework in the Age of Artificial Intelligence: From Agenda-setting to Policy Evaluation,» *Government Information Quarterly*, vol. 37, nº 4, 2020.
- [10] Deloitte, «¿Qué es la Industria 4.0?,» <https://www2.deloitte.com/es/es/pages/manufacturing/articles/que-es-la-industria-4.0.html>.
- [11] M. Guardamagna y M. Reyes, «El Desafío de la Implementación de Políticas Públicas Participativas para el Desarrollo del Territorio,» *Economía, sociedad y territorio*, vol. 19, nº 59, pp. 1003–1033, 2019.
- [12] C. d. P. C. y. C. Social, «Políticas Públicas de Participación Ciudadana, Control Social, Rendición de Cuentas, Transparencia y Lucha contra la Corrupción,» 2014. <https://www.cpcs.gov.ec/wp-content/uploads/2016/02/POLITICAS-PUBLICAS.pdf>.

Pautas Lógicas para el Diseño Computacional de Sistemas Críticos Basadas en el Paradigma de Fitch y Muckin

Miguel Torrealba
mtorrealba@usb.ve

Departamento de Computación y Tecnología de la Información, Universidad Simón Bolívar, Sartenejas, Venezuela

Resumen: Los sistemas tecnológicos críticos del ciberespacio soportan servicios y operaciones que inciden agudamente en el funcionamiento regular de nuestras sociedades modernas. Además, desde décadas atrás su buena operatividad está en la mira de los conflictos políticos y bélicos de hoy. En consecuencia, su defensa resulta primordial y demanda una elaboración de un diseño robusto, privilegiando el rol de la ingeniería de seguridad en conjunto con el de la ingeniería de software. Las arquitecturas de sistemas que se defienden por sí mismas, son una aproximación alterna al esquema tradicional de asegurar el sistema conforme a la identificación y tratamiento de los servicios de seguridad. Este trabajo expone un conjunto de pautas organizadas, para guiar la construcción de un diseño de arquitecturas críticas con capacidad de defensa, que use el paradigma de Fitch y Muckin y deje espacio para aplicar variaciones en la implementación y ejecución del mismo.

Palabras Claves: Ingeniería de la Seguridad; Ciberseguridad; Seguridad por Diseño; Defensa Inteligente; Sistemas Críticos; Ciberguerra; Ciberconflicto.

1. INTRODUCCIÓN

En 2019 Scott C. Fitch y Michael Muckin, en la empresa estadounidense Lockheed Martin, publicaron un artículo de su trabajo corporativo (“White Paper”) con nombre: “*Defendable Architectures. Achieving Cyber Security by Designing for Intelligence Driven Defense*” [1]. Ese trabajo sostiene una crítica al diseño tradicional de sistemas informáticos, en materia de ciberseguridad; ellos señalaron que es muy ambicioso aplicar la orientación de diseñar bajo la pauta de protección con controles, por anticipado sobre los sistemas y esperar, que en su posterior ejecución, estos se sostengan ante cualquier ataque. En su lugar propusieron una aproximación de diseño que se soporta sobre la recopilación y procesamiento de elementos de inteligencia; insumos que trabajan con riesgos informáticos que permitan derivar la arquitectura de un sistema, capaz de defenderse.

De manera que la orientación alterna de Fitch y Muckin, es un recorrido de procesos que inicia con el diseño, para luego ejecutar etapas de construcción, operación y llegar al momento de la defensa. Esta visión posee el atractivo de que ante la diversidad y variabilidad de los ataques, la resiliencia y protección que se obtendrá no deberá ser tan genérica como actualmente ocurre y además, será más adaptativa ante los peligros reales que sucedan. Una propuesta distinta, si se toma en cuenta la complejidad del área a tratar y los decepcionantes resultados que hoy en día se observan.

Por otra parte, actualmente el diseño de Sistemas e Infraestructuras Críticas, aún sigue el esquema ortodoxo y común de cualquier sistema informático, de modo que ante escenarios de Ciberconflictos [2] y Ciberguerras [3], ello representa un problema real para nuestro mundo moderno, basado en tecnologías digitales. La exposición y vulnerabilidad de estos sistemas, tan sensibles, como son los militares, los industriales, de servicios básicos -electricidad, agua,

comunicaciones- y soporte para operaciones diarias de la banca o comercio, resulta bastante parecida a la de un sistema informático o gerencial común. Algo que se ha agudizado al añadir características de atención al público e interfaces para el usuario, con funcionalidad del World Wide Web.

La aproximación más común que aún se aplica sobre el diseño de nuestra tecnología digital, se basa en la incorporación de técnicas e instrumentos de protección muy asociados con los clásicos servicios de seguridad de la norma X.800 del ITU-T; confidencialidad, integridad y disponibilidad. Bajo esa línea, lo que la ingeniería de sistemas hace es parecido a crear un artefacto y luego colocar protecciones, inhibidores o elementos para disuadir los ataques. El empleo de la criptografía y los distintos sistemas de prevención, filtrado y/o desviación, encajan en esa orientación. Ello resulta válido, aunque la realidad de lo numeroso de los sistemas comprometidos, o los dañados, inoperantes y también de los alterados indebidamente, algunos por años consecutivos sin que se advierta su irregular condición, es un signo más de que esta aproximación no siempre es ideal como pauta de diseño.

Este trabajo propone un conjunto de pautas lógicas para profundizar en la idea de Fitch y Muckin, y así diseñar arquitecturas de sistemas críticos, que soporten mejor los ataques externos y faciliten, inteligentemente, la recuperación y restauración operativa de los mismos.

2. SISTEMAS CRÍTICOS Y LOS TEATROS DE CIBERCONFLICTO Y CIBERGUERRA

En nuestro mundo moderno, situaciones de Ciberconflicto o Ciberguerra se vienen haciendo más comunes y los ataques sobre infraestructuras y sistemas cibernéticos, se vislumbran como un riesgo factible. Varias naciones ya han establecido componentes militares para el control del ciberespacio, así como han emitido guías para construir y mejorar su

infraestructura tecnológica con soporte a la Internet o los servicios con la Nube.

Así por ejemplo, el Departamento de Defensa de los EEUU (DoD) percibe al ciberespacio como un Dominio Global y en 2011 declaró, lo siguiente como parte de su estrategia de defensa del Ciberespacio: *“El DoD usa el ciberespacio para habilitar sus operaciones militares, de inteligencia y negocios, incluyendo los movimientos de personal y materiales, así como el comando y control del espectro completo de sus operaciones militares”*.

Por otra parte, en las últimas décadas, los incidentes directos e indirectos en todo el mundo, sobre plantas nucleares, industriales, eléctricas, aeropuertos y hospitales, han crecido. Algunos se han derivado de fallas básicas como es la actualización de una protección del tipo antivirus, incrustado en un Sistema Operativo [4]. Mientras que otras fueron dirigidas y específicas a un blanco, como es el caso de Stuxnet y los sistemas con tecnología centrífuga de la planta de desarrollo nuclear de Irán [5].

En 2020 los investigadores Sungjoong, Jiwon, Haengrok, Dongil y Dongkyoo sostuvieron en un trabajo de IEEE Access que: *“La ciberguerra ocurre en el ciberespacio y está siendo procesada y generada a través de varios ciber ataques, tales como el hackeo de sistemas de información militar de otros países y paralizando sistemas militares y de información para la defensa que permiten alcanzar objetivos militares”* [6]. Este panorama delicado, no son los “hacking” de finales de los noventa, sino que también incorpora naciones – estados atacando y defendiendo su ciberespacio y vínculo con su soporte al espacio real, lo cual obliga a que los profesionales de computación se replanteen preguntas básicas.

Dudas sobre la inseguridad que verdaderamente rodea los desarrollos y artefactos tecnológicos de soporte crítico, para servicios básicos, sistemas de producción y economía moderna. Consultas como: ¿el esquema para desarrollar un sistema y luego colocar controles y protecciones, es suficientemente bueno y sólido para las necesidades?. Adicionalmente, si se diseñan sistemas teniendo presente las necesidades y requerimientos de seguridad desde su inicio, ¿verdaderamente se puede anticipar todos los peligros y cambios de contexto que un mundo exterior puede producir?.

Basta con mirar la variedad y sofisticación de los ataques informáticos de nuestros días, así como las cifras de pérdidas económicas por fallas de ciberseguridad, para vislumbrar la respuesta del caso. Finalmente, también hay espacio para la cuestión central de esta reflexión: ¿el resguardo de un sistema crítico debe apoyarse en el mismo tipo de instrumentos y técnicas de seguridad, que se incorporan en un sistema informático común, o amerita un modo diferente y más profundo de concepción?

3. TECNOLOGÍA CRÍTICA CON BASE AL PENSAMIENTO DE FITCH Y MUCKIN

Tomando la idea central de las arquitecturas que se defienden, se puede enumerar los siguientes aspectos que deben guiar el trabajo de diseño de los sistemas críticos:

- Los sistemas no siempre operan en las mismas condiciones, su contexto cambia y ello potencialmente incide sobre la inseguridad de fuente externa. En consecuencia, sus respuestas funcionales no siempre deben ser las mismas.
- La exposición y visibilidad del sistema ante el mundo exterior, debe poder ser reducida si así se requiere.
- La concepción de un sistema crítico se debe soportar en la aplicación de una inteligencia adaptativa, que pueda responder ante la presencia y comportamiento de cualquier ataque.
- Las protecciones deben responder a las técnicas de los atacantes, al igual que a las nuevas amenazas. Por ser sistemas de naturaleza crítica, la guía X.800 ITU-T de los servicios de seguridad no es el elemento determinante. Es un complemento para enfocar.
- El diseño interno debe ofrecer variaciones en su funcionalidad, haciendo progresiva la fortaleza de su defensa. De ser necesario, debe presentar, selectivamente, una representación falsa del sistema.
- La inter-relación humana es indispensable y comprende las acciones agresivas. Como los ataques se responden en menos que segundos, las máquinas son ideales para eso, pero detectar si lo que ocurre es un falso positivo, una distracción o es la esencia de un verdadero mal, es algo que los humanos entrenados aún hacen mejor. La decisión de atacar debe privilegiar lo humano.
- En semejanza con los seres humanos, ante un peligro inminente, la última defensa es “correr o pelear”. Por ello el sistema crítico deberá considerar la disyuntiva, no excluyente, entre migrar datos, operaciones y/o agredir.

Por otra parte, existen aspectos que son ejes de soporte en la seguridad final del diseño y deberán ser atendidas con métodos no necesariamente clásicos. El primero de ellos se vincula con las relaciones y dependencias basadas en la confianza del sistema.

4. EL DISEÑO DE SISTEMAS CRÍTICOS Y LA CONFIANZA

Un elemento relevante en el diseño de los sistemas críticos es la confianza y ello incide en momentos de peligro. Los sistemas tradicionales se elaboran teniendo en mente una confianza estática. Este es el caso de una actualización de software del fabricante que se dispara automáticamente, el de una cuenta de usuario privilegiada que todo lo puede hacer, o tal vez el de una autenticación única de usuario o aplicación (“Single Sign-On”), que provee acceso a todo un ambiente de red. Otro caso es el de un rango de direcciones de redes predefinidas y reconocidas como válidas, o segmentos de tráfico en las comunicaciones que de acuerdo al perímetro lógico se dan como “confiables” o incluso, pudiera ser el de conexiones de servicio y/o mantenimiento “amigables”, que fabricantes y/o proveedores actualmente incluyen en sus contratos. El problema es que si uno de esos elementos se compromete, tradicionalmente, las protecciones no lo reconocen y no pueden responder apropiadamente. Se sigue considerando que son fiables y pueden ser mal aprovechadas.

De forma que esa aproximación tan comúnmente usada, puede ser contraproducente para un sistema crítico, ya que a menudo los atacantes explotan las debilidades de confianza y con ello penetran o extienden su ámbito de acción. Luego, es posible suponer que la confianza debe variar y en situaciones de riesgo o ataque, reducirse sustancialmente para facilitar una mejor protección. La noción fundamental es que la confianza sostenga la ejecución de las operaciones e intercambios y de ser requerido, pueda alterar la funcionalidad “normal” del sistema crítico.

La Figura 1 representa, sobre un sistema crítico, un esquema simple para manejar diferentes consideraciones dinámicas de la confianza.

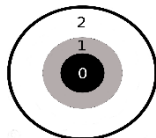


Figura 1: Esquema Jerárquico de la Confianza

A través de tres círculos concéntricos, que representan la extensión de la confianza, se puede apreciar cuál conjunto contiene a otro. Así pues, de adentro para afuera está el nivel “0” de confianza, que es mínimo y se visualiza como el núcleo. Ese ámbito refleja las Relaciones de Confianza Básicas y comúnmente refiere la de administradores y operadores locales, así como aplicaciones propias, no en red. Se espera también, que este nivel predomine en situaciones de ataques, donde hay que disminuir las suposiciones y creencias. El nivel “1”, que se ve como intermedio, amplía la confianza a un área de Relaciones de Confianza Regionales, es decir, donde se puede aceptar que las redes locales y/o dominios corporativos son “amigables”. El ámbito externo queda afuera de este nivel, por lo cuál ha sido concebido para ejecuciones de riesgo moderado o de sospecha. El último nivel, identificado por “2”, que se ve como el círculo más externo, incluye objetos y asociaciones de confianza con el exterior del sistema. En situaciones de tranquilidad pudiera llegar a fronteras con los proveedores conocidos de suministros, servicios y/o comunicaciones, que interconectan con la Red de Redes.

Colateralmente, la decisión de cuál nivel de confianza debe ser el vigente puede permitirse que sea automatizada, bajo supervisión de indicadores, métricas y perfiles históricos, siendo lo común que se escale gradualmente, pero debe aceptarse la intervención explícita y humana para fijarla arbitrariamente. Una manipulación, que por ninguna razón debe ser remota y bien puede exigir doble autorización de entes diferentes, en ubicaciones separadas y con factor de autenticación triple.

5. EL ESTADO Y EL DISEÑO DE SISTEMAS CRÍTICOS

Otro elemento de importancia que resulta ser una columna sobre la que se edifica la seguridad de un sistema crítico, es su condición operacional. A eso y a la inteligencia recabada de señales de riesgos de su contexto, denominamos estado. El estado define la “salud y situación” del sistema crítico y en

consecuencia, su capacidad funcional y la atención ante el exterior. Un sistema bajo ataque deberá tener un estado que como tal lo represente y activar o reforzar sus medidas de salvaguarda o respuesta, ante tal condición.

Un modelo simple de estado incluye cuatro categorías: normal, atípico, en-riesgo y bajo-ataque.

- Normal: El sistema opera con sus *relaciones de confianza* pre-establecidas y sus medidas de protección tradicionales. En principio, se puede esperar que este estado se corresponda con un nivel de confianza “2”.
- A-típico: emite alertas y eleva los niveles de supervisión. Hay observación y comparación de las relaciones de confianza contra su perfil tradicional. Se refuerza el uso de indicadores y métricas estadísticas. Bien puede acoplarse con los niveles “1” y “2”. También emplea técnicas de Inteligencia Artificial para gestionar los riesgos presentes.
- En-riesgo: eleva considerablemente los niveles de protección, control y supervisión; internamente registra todo la actividad que puede y no atiende nuevas peticiones, que no sean las tradicionales. Emite alertas sobre todos los niveles y rechaza relaciones de confianza indirectas. Opera con la salvaguarda de “bajo extrema emergencia” se estacionará en un “área de resguardo temporal”. Se asocia con los niveles de confianza “0” o “1”.
- Bajo-ataque: No atiende nuevas peticiones, hay selectividad en lo que ya se realiza y se limita el flujo de información. Se elevan los controles y protecciones a su máxima capacidad. Se bloquean servicios no críticos y se inicia procedimientos internos de contingencia (incremento de alertas, comunicación de su condición a sistemas con los que se colabora, transferencia de funcionalidades, activación del respaldo y se puede activar la migración de datos sensibles). Si no hay respuesta estratégica, sigue tácticas disuasivas o de confusión hacia el exterior. Es fundamental que se registre todo lo que se pueda afuera de sí mismo y si no se obtiene apoyo, progresivamente el sistema se desactiva. Se reduce la confianza a su nivel núcleo, “0”.

Al contrario de lo que a menudo los diseñadores presumen, hay que considerar situaciones de final no feliz. Posibilidad de que el sistema no posea las defensas adecuadas para manejar algún ataque. A eso lo llamamos “situaciones extremas”.

6. SITUACIONES EXTREMAS Y LAS DIRECTIVAS HUMANAS

Una diferencia notable entre un sistema informático y un sistema crítico, deberá ser su ajuste en la estrategia de defensa que se aplique. Desde un nivel estrictamente defensivo a otro que incluya acciones ofensivas; algo propio de ciberconflictos y/o ciberguerras. Obviamente, este giro debe contar con la autorización humana y nuevamente, puede ser necesario que esta sea dual e incorpore tecnologías biométricas.

La Figura 2 describe una organización lógica de una potencial arquitectura, reflejando los perímetros de defensas y el externo al sistema. También presenta un espacio de instrumentos para soportar pasivamente o para hacerlo, con acciones agresivas.

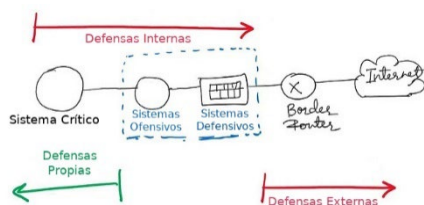


Figura 2: Arquitectura con Defensas y Armas

Esa gráfica expone áreas que rodean al sistema crítico e ilustra zonas para defender. Hay defensas externas que inician en el enlace con los proveedores de comunicaciones y servicios con la Internet (ISP). Otras defensas son internas y graduales, llegando hasta ser propias y particulares del mismo sistema crítico. Entre ambas áreas existen entidades para protecciones defensivas y otras hostiles. Estas últimas se emplean en situaciones de “supervivencia” del sistema crítico y aplican en caso de ataques, que saturan o sobrecargan, como las del tipo Negación de Servicio Distribuido (DDoS).

Estos escenarios pueden ser el último recurso para evitar apagar al sistema como protección final. Llegar a una desactivación tan drástica, es la última defensa y únicamente tiene por objetivo preservar al sistema para otra ocasión, pero su razón de ser podría haberse perdido. Entonces resulta lógico esperar que este tipo de situaciones no sean propias de ambientes tradicionales y se acomodan, a otras más cercanas a la guerra. Las valoraciones y decisiones a seguir en esos conflictos escapan de la construcción técnica, pero deben poder ser instrumentadas con apoyo tecnológico.

7. CONTROLES, PROCEDIMIENTOS Y PRINCIPIOS DE SEGURIDAD

Para completar la etapa de diseño se debe poder estipular los procedimientos que permitirán tratar con la capacidad de inteligencia de la seguridad del sistema. Para ello continuamente habrá que adaptar el Modelo de Amenazas, lo cual constituirá un proceso propio del sistema crítico, relacionado con su propia seguridad. En otras palabras, esto se constituye como un mecanismo de cuidado del mismo y sobrepasa la simple activación de un control o una protección. Un procedimiento que incorpora un nivel de inteligencia para cambiar de forma según se demanda, a modo de morfogénesis.

Resulta necesario agregar que los procesos y procedimientos de seguridad, deberán estar alineados con aquellos principios de seguridad de Saltzer y Schroeder [7] que sean escogidos. Entonces, en función de los resultados se pueden establecer los controles, las protecciones, los modelos y las técnicas que soportarán todos esos mecanismos. Cada uno de estos son diferentes pero se relacionan, armoniosamente, entre sí. Por ejemplo, un modelo para establecer la aleatoriedad de una

“función hash” o de un mecanismo criptográfico, puede ser el modelo de seguridad del oráculo [8]. Mientras que la herramienta que instrumenta la función matemática señalada, es un instrumento que puede ser empleado por un procedimiento, para determinar si dos secuencias de bits son iguales. Con ello se puede tener la facultad de determinar si un archivo, por ejemplo, fue alterado en su contenido o si se corresponde con lo que antes se transmitió en forma protegida.

Por otro lado, una técnica de seguridad es por ejemplo el uso de señuelos o distracciones. Este tipo de elementos pueden ser usados para construir medidas evasivas o ganar tiempo y se deben acomodar, dentro de procedimientos de defensa. Cuando y cómo habilitarlos dependerá de la inteligencia recabada y la evaluación de riesgos, que previamente se haya realizado. Algo que puede ejecutarse en tiempo real y así variar, según el peligro presente. Bajo esta aproximación, es posible deducir que el diseño que se realice puede obtener libertad o mucha flexibilidad, para los detalles de la futura implementación del mismo sistema. Aún cuando, hay una línea directriz suprema que regula y dirige la protección del sistema.

8. COLOFÓN

Este trabajo ha descrito una serie de pautas lógicas que pueden guiar el diseño de seguridad de un sistema crítico, que además se apoya en la propuesta de Fitch y Muckin para sistemas con capacidad inteligente de defensa. Esto aspira a ser una alternativa al tradicional esquema de diseño de sistemas y luego colocar las protecciones, según se identifiquen los servicios de seguridad X.800 del ITU-T.

REFERENCIAS

- [1] S. Fitch y M. Muckin, *Defendable Architectures Achieving Cyber Security by Designing for Intelligence Driven Defense*, LM White Paper, PDF, 2019. <https://surl.li/hsjrbq>
- [2] N. Petru-Cristian, *Cyber Conflict and International Relations: A Comprehensive Analysis of Cyber Deterrence Strategies en Contemporary Geopolitics*. 10.13140/RG.2.2.29742.69449, 2024
- [3] T. Berson y D. Dorothy, *Cyberwarfare*. IEEE Security & Privacy. 9. 13-15. 10.1109/MSP.2011.132, 2011.
- [4] HKCERT, *CrowdStrike Denial of Service Alert*, HTML, Julio 2024. <https://shorturl.at/2Vvka>
- [5] R. Langner, *To Kill a Centrifuge. A Technical Analysis of What Stuxnet's Creators Tried to Achieve*, PDF, Noviembre 2013. <https://shorturl.at/TVKle>
- [6] S. Kim, J. Kang, H. Oh, D. Shin y D. Shin, *Operation Framework Including Cyber Warfare Execution Process and Operational Concepts*, IEEE Access, vol. 8, pp. 109168-109176, doi: 10.1109/ACCESS.2020.3001286. 2020
- [7] J. Saltzer y M. Schroeder. *The Protection of Information in Computer Systems*, HTML.1975. <http://www.cs.virginia.edu/~evans/cs551/saltzer/>
- [8] M. Bellare y P. Rogaway. *Random Oracles are Practical: A Paradigm for Designing Efficient Protocols*. Proceedings of the First ACM Conference on Computer and Communications Security, 1993.

Proceso de Análisis del Dominio para Arquitecturas de Referencias IMFS & VCD Modelado con SPEM

Aleidys Arraiz¹, Yusneyi Carballo-Barrera²
aleidysarraiz@gmail.com, yusneyicarballo@gmail.com

¹ Departamento de Informática, Universidad Politécnica Territorial del Estado Aragua, La Victoria, Venezuela

² Escuela de Computación, Universidad Central de Venezuela, Caracas, Venezuela

Resumen: Para construir arquitecturas de referencias en el ámbito de la industrialización de software se utilizaron los elementos planteados por el enfoque de Líneas de Productos de Software (LPS) integrados en Fábricas de Software como primer enfoque central, que en su etapa de Ingeniería de Dominio (ID), contempla dos disciplinas: Análisis del Dominio y Diseño del Dominio. La primera, permite definir y estructurar los requisitos del dominio que han sido elicitados previamente, para garantizar la genericidad de la Arquitectura de Referencia o plantilla genérica de donde se derivan productos concretos de la familia de sistemas de software. La segunda, es la disciplina en donde se construye la Arquitectura de Referencia, es la realización del dominio donde se completa la información sobre los componentes que puedan ser instanciados o puntos de variabilidad, además de sus instancias concretas o variantes. En este trabajo se presenta la especificación de la disciplina de Análisis del Dominio, adaptada al proceso de transformaciones derivadas de la Ingeniería de Modelos (IM) como segundo enfoque central. También se describe cómo las transformaciones generan modelos que favorecen la construcción de una Arquitectura de Referencia en el Esquema de Fábricas de Software (EFS) utilizando un enfoque de Calidad exclusivo o específico para el dominio, tal es el caso de la Vista de Calidad del Dominio (VCD). Como un aporte importante de este trabajo, se realiza una revisión del estándar de Modelo de Calidad del Producto ISO/IEC 25010:2023. La especificación de las disciplinas se expresa en términos de representación de modelos de procesos de ingeniería de software a través del estándar SPEM 2.0 para un modelo un proceso IMFS & VCD (Ingeniería de Modelos, Fábricas de Software y Vista de Calidad del Dominio).

Palabras clave: Fábricas de Software; Arquitectura de Referencia IMFS & VCD; Ingeniería del Dominio; Ingeniería de Modelos; SPEM; Calidad de Software; ISO/IEC 25010:2023.

1. INTRODUCCIÓN

IMFS & VCD es un modelo de proceso con enfoque *top-down* de descripción arquitectural que contempla dos fases primordiales de acuerdo con las Líneas de Productos de Software (LPS): Análisis del Dominio y Diseño del Dominio. Este modelo involucra a la Ingeniería de Modelos, las Fábricas de Software y la Vista de Calidad del Dominio. Está basado en el Esquema, uno de los elementos más relevantes en el desarrollo de software bajo el enfoque de Fábricas de Software. El esquema define los puntos de vistas que son necesarios para la construcción de un tipo de producto de software o miembro perteneciente a una familia de sistemas de software.

En detalle, una vista es una representación de todo el sistema de software desde una determinada perspectiva establecida por los involucrados en el desarrollo de la familia de sistemas de software [1]. Cada uno de estos puntos de vista son el inicio para identificar el núcleo (en inglés *core*) de artefactos que permitirá el camino más eficiente para producir software porque contiene los mecanismos para expresar las vistas. En otras palabras, su despliegue depende de los componentes estructurales proporcionados por las vistas en el Esquema de Fábricas de Software, que posteriormente se implementan en la Plantilla de acuerdo con el modelo conceptual de Fábricas de Software, que no es más que la colección de todos los activos definidos en él. Adicionalmente IMFS & VCD se fundamenta en la perspectiva de calidad de software del dominio planteada por Losavio y

Matteo en 2013 [2] e incluye las técnicas para transformaciones de modelos del enfoque de Ingeniería de Modelos guiadas a través del estándar MDA (*Model Driven Architecture*). Por otra parte, en el contexto de desarrollo industrial de software, en el que los procesos se visualizan como productos por su constante cambio y evolución, los procesos deben ser configurados para que puedan ser adaptados a las necesidades del dominio. Se hace necesario entonces, la definición de un lenguaje que permita la construcción de los productos de las familias de sistemas de software de forma adecuada, lo que implica un proceso complejo, ya que contempla múltiples actividades que se trabajan de forma independiente y abarcan etapas como la gestión, los aspectos de calidad y el soporte. El modelado de procesos de software es la técnica que a través de la cual se optimiza el desarrollo y calidad de los productos de software resultantes [3].

La Ingeniería de Procesos de Software (SPE o *Software Process Engineering*) se utiliza para modelar, diseñar y optimizar procesos haciendo uso de lenguajes para el modelado de procesos existentes que se fundamentan en flujos de trabajo. Uno de los lenguajes utilizados con este fin es SPEM (*Software Process Engineering Metamodel*) de la OMG [4]. Este lenguaje está basado en MOF (*MetaObject Facility*) el cual forma parte del Lenguaje Unificado de Modelado o *Unified Modeling Language*, por lo que la notación de SPEM está basada el estándar UML. Además, soporta la representación de familias de sistemas de software con sus componentes, considerando un

tipo de ontología para procesos de desarrollo de software con elementos de modelado de procesos que permite describirlos sin ningún tipo de restricción en el aspecto de las disciplinas que los componen.

SPEM provee una sintaxis y una estructura para cada etapa del desarrollo las cuales pueden ser aplicadas en familias de productos de software, porque contempla: roles, tareas, artefactos, lista de verificación, productos de trabajo, técnicas y herramientas, estructuras de trabajo, capacidad de rastreo y refinamiento, ayuda sensible al contexto, guía y lineamientos, así como, descripción textual de elementos [5]. SPEM es utilizado para modelar los procesos de Análisis de Dominio en el contexto de Ingeniería de Modelos y Fábricas de Software (IMFS) con aspectos de Vistas de Calidad de Dominio (VCD). Al ser un estándar basado en UML, SPEM es independiente de las metodologías que se utilicen para el desarrollo de software y ofrece una visión detallada en un nivel de abstracción alto, donde se evidencian los elementos fundamentales en un proceso que se encuentra constituido por roles, actividades, tareas y artefactos, entre otros, incluyendo también la relación y trazabilidad que guardan entre sí.

En este trabajo se presenta la especificación del proceso correspondiente a la disciplina de Análisis del Dominio del modelo, la cual ha sido denominada IMFS & VCD [3], con una introducción a la actualización del estándar más reciente basado en el modelo de calidad del producto ISO/IEC 25010:2023 [6], y su representación en SPEM. La finalidad es dar a conocer la disciplina de Análisis del Dominio del modelo de proceso, resaltando sus entradas, salidas, roles responsables, técnicas, herramientas y estándares utilizados, así como los enfoques que han sido empleados en su construcción, mostrando una primera parte de su trazabilidad y sus componentes.

2. ANÁLISIS DE ENFOQUES, MÉTODOS Y TÉCNICAS PARA IMFS & VCD

Con el estudio exhaustivo de antecedentes y los artículos de investigación derivados de la tesis doctoral de Arraiz en 2024 [3] se puede afirmar que, aunque hay investigaciones que integran el enfoque de Fábricas de Software con la Ingeniería de Modelos, ninguno había definido un proceso proactivo *top-down* que estudie el dominio de familias de productos de software como el que se propone, proveyendo todas las fases, disciplinas, artefactos, roles, tareas, que permitan derivar una Arquitectura de Referencia para un dominio de familia de sistemas de software en particular.

Es por ello que, se determinó como primer paso, la relación de los enfoques de Fábricas de Software con Ingeniería de Modelos presentados en diversas publicaciones científicas que utilizan el modelado de negocio para la caracterización del dominio. Se incluye la visión presentada por Ariste, Ponisio, Nahuel y Giandini en 2007 [7] que abarca el diseño transformaciones de modelos CIM/PIM (*Computational Independent Model / Platform Independent Model*), desde un enfoque de negocio hacia un enfoque de sistema, cumpliendo con la premisa de incluir aspectos de calidad en el proceso de Análisis del Dominio de la disciplina de Ingeniería del Dominio contenida en el Esquema de Fábricas de Software. Además, se integra el enfoque de Vista de Calidad del Dominio (VCD) [8], basada en el estándar ISO/IEC 25010, incorporando la actualización

publicada en 2023 por la ISO [6] para las actividades de la disciplina de Análisis del Dominio.

Una vez sustentada la VCD con las Fábricas de Software, se considera el trabajo realizado por Arraiz, Losavio y Matteo en 2015 donde presenta una revisión sistemática de trabajos basada en estos enfoques [9]. Luego, se tiene el trabajo de Arraiz y Esteller publicado en 2018 [10], en el que se describen las etapas para la construcción del modelo IMFS & VCD, por lo que se incluyen los elementos de transformaciones de modelos en el Esquema de las Fábricas de Software definido en un modelo conceptual de dominio. Lo anterior constituye la base para la derivación de una Arquitectura de Referencia de familias de sistemas de software, partiendo desde un modelo CIM (*Computational Independent Model*) hacia un modelo PIM (*Platform Independent Model*), que considera la trazabilidad de requisitos y de artefactos, junto a la experticia de personas profesionales en el dominio.

Otro punto a considerar, es que el modelo de proceso que se plantea se describe en el lenguaje de modelado de procesos de ingeniería de software SPEM 2.0. La selección de este lenguaje se debe a que soporta la representación de familias de sistemas de software en un dominio determinado en el contexto de desarrollo industrial de software. En la Figura 1, se presenta el modelo conceptual de Fábricas de Software con Ingeniería de Modelo, que se establece como modelo de dominio base para el desarrollo del proceso que se ha construido y que representó uno de los primeros resultados obtenidos en los trabajos que antecedieron a esta investigación.

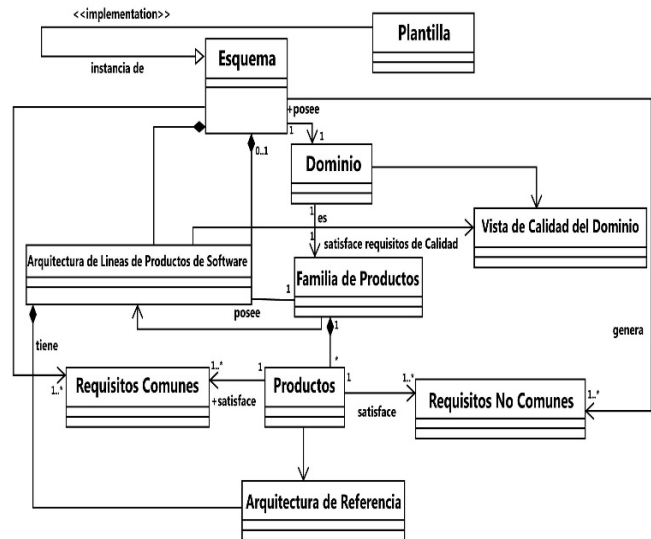


Figura 1: Modelo Conceptual de Fábricas de Software con VCD para las Transformaciones

La integración de los enfoques, métodos, técnicas para el proceso de Análisis del Dominio se encuentra representada para IMFS & VCD, como se observa en la Figura 2. Como enfoque de calidad el proceso IMFS & VCD, se ha empleado la versión de ISO/IEC 25010 publicada en 2011. Sin embargo, en este trabajo se realiza una consideración basada en la introducción para la actualización ISO/IEC 25010:2023 [6], en la que se ha incluido una característica de calidad adicional al modelo de

calidad del producto la cual han denominada flexibilidad, sustituyendo el concepto de portabilidad.

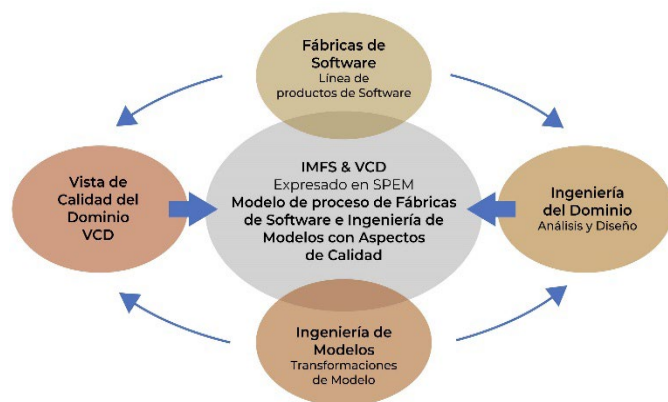


Figura 2: Integración de Enfoques para el Modelo IMFS & VCD

También se han incluido subcaracterísticas nuevas a la categoría de usabilidad, la cual en la actualización ISO/IEC 25010:2023 es identificada como Capacidad de Interacción, ampliando los alcances de sus dimensiones al incluir ahora la participación del usuario, la asistencia al usuario por parte del software y, la protección contra errores de usuario, como puede observarse en la Tabla I.

Tabla I: Modelo de Calidad del Producto ISO/IEC 25010:2023

Adecuación Funcional	Eficiencia	Compatibilidad	Capacidad de Interacción	Confiabilidad	Seguridad	Mantenibilidad	Flexibilidad
Complejidad	Comportamiento en Tiempo	Coexistencia	Reconocibilidad de adecuación	Madurez	Confidencialidad	Modularidad	Adaptabilidad
Correctitud	Uso de Recursos	Interoperabilidad	Capacidad de ser aprendido	Disponibilidad	Integridad	Reutilización	Capacidad de ser instalado
Capacidad de ser apropiado	Capacidad		Capacidad de ser operado	Tolerancia a fallas	Capacidad de no ser rechazado	Capacidad de ser analizado	Capacidad de ser reemplazado
			Protección contra errores de usuario	Capacidad de recuperarse	Capacidad de ser auditado	Capacidad de ser modificado	Capacidad de adaptabilidad
			Participación del usuario		Autenticidad	Capacidad de ser probado	
			Inclusividad				
			Asistencia al usuario				
			Capacidad de ser autodescriptivo				

El proceso de análisis del dominio de IMFS & VCD, que a continuación se describe, presenta la descripción de entradas y salidas de esta disciplina, para posteriormente incluir representaciones gráficas de las actividades en modelos de procesos de ingeniería de software a través del estándar SPEM 2.0. Se suministra a continuación la descripción entradas y salidas de cada actividad.

Fase I: Análisis del Dominio

Actividad 1. Estudio y Conocimiento del Dominio

Entradas:

- Descripción de forma textual y gráfica del problema en el contexto de la familia de sistemas de software, para la elaboración de un modelo de negocios específico del dominio [11].

Se debe especificar los procesos inherentes al negocio siguiendo lineamientos de la Gestión de Procesos de

Negocios de una organización con la finalidad de identificar y comprender la información del dominio, especificando los detalles de los Requisitos Tecnológicos y estableciendo las Reglas de Negocio (RN), lo que permite un análisis, modelado y asociación con los RF y los RNF. Los procesos a los cuales se hace referencia son:

- Descubrimiento.
- Análisis.
- Desarrollo.
- Monitoreo.
- Optimización.

Salidas:

- Modelo de Negocio, utilizando la notación de modelado de procesos de negocio (BPMN) [11]. El modelo de Negocio se considera como el modelo origen para iniciar el proceso de transformación utilizando la visión del enfoque de Ingeniería de Modelos.
- Los Requisitos Funcionales (RF) y los Requisitos No Funcionales (RNF) del modelo de calidad, ahora en el estándar ISO/IEC 25010:2023 [6].
- Las Reglas de Negocio (RN) del dominio perteneciente a la familia de sistemas de software.

Actividad 2. Descripción para las Transformaciones de Modelos

Entradas:

- Modelo de negocio especificado en BPMN (Modelo Origen).
- Reglas de Negocio (RN).
- Requisitos Funcionales (RF) y Requisitos No Funcionales (RNF) del dominio con base en el ISO/IEC 25010.

En esta actividad se realiza la identificación de las Reglas de Transformaciones (RT) gráficas de BPMN a UML que se utilizan para el subproceso de transformaciones, considerando que ambos lenguajes son de notación gráfica lo cual permite definir de una forma clara las Reglas de Transformación (RT). También se realiza la descripción de las RT, tomando en consideración las siguientes premisas para establecer la correspondencia entre el modelo origen y el modelo destino, a fin de pasar del nivel CIM al PIM con transformaciones de tipo M2M:

- **P1.** Las RT involucran un artefacto de metamodelo BPMN y uno o más artefactos del metamodelo UML.
- **P2.** Se consideran: transformaciones 1 a 1, es decir, de elemento simple de BPMN a elemento simple de UML, y también, transformaciones 1 a N, es decir, de elemento simple de BPMN a elemento compuesto o conjunto de elementos compuestos de UML.
- **P3.** Se utiliza el metamodelo correspondiente a los diagramas de Casos de Uso UML, ya que éste representa el modelo de comportamiento, en sincronía con el comportamiento que establecen los diagramas de procesos de negocio en BPMN.

Para establecer las transformaciones en el modelo de proceso

IMFS & VCD, se consideró la Notación de metamodelo para Diagramas de Procesos de Negocio (DPN) de BPMN 2.0 de la OMG (2013), que se presentan en la Figura 3.

Se realiza la construcción de tabla contentiva de especificaciones gráficas de los diagramas de procesos de negocio en BPMN hacia los diagramas de casos de uso UML y Construir el Modelo de casos de uso resultante (Modelo Destino) con todos los diagramas de casos de uso correspondientes a la descripción de dominio de la Familia de Sistemas de Software estudiado, una vez que se hayan aplicado las reglas de transformación.

Salidas: modelos de Casos de Uso (Modelo Destino), resultante de la aplicación de las RT.

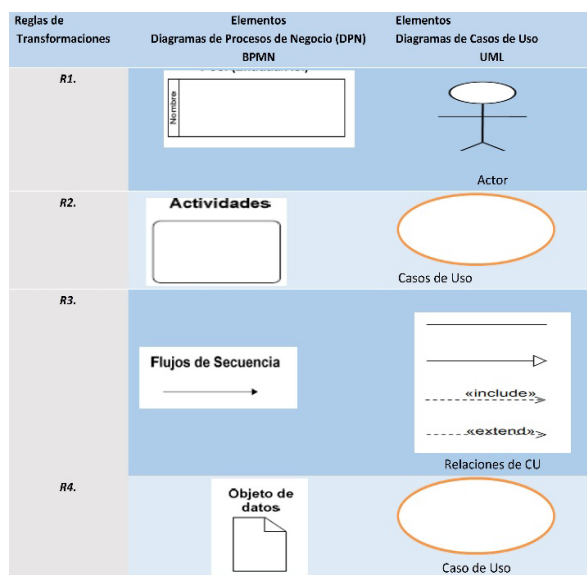


Figura 3: Especificación Gráfica de Correspondencia entre los Diagramas de Procesos de Negocio en BPMN y los Diagramas de Casos de Uso de UML

Actividad 3. Descripción de RF y RNF para Establecer el Modelo VCD de la Familia de Sistemas de Software

Entradas:

- Modelo de Casos de Uso (resultante de la actividad 2).

En la actividad se deben definir los activos funcionales del dominio: especificar los principales RF comunes del dominio de la familia de Sistemas de Software, especificar los RNF, Señalando y describiendo detalladamente los RNF que influyen de acuerdo a la familia de sistemas de software estudiada, documentar los RF y RNF a través del ERS (Especificación de Requisitos de Software), documento para la especificación basado en el estándar IEEE-830-1998 [11].

Salidas:

- ERS de la familia de sistemas de software [12].

Actividad 4. Especificación de Vista de Calidad Funcional del Dominio (VCFD)

Entradas:

- ERS de la familia de sistemas de software [12].

Se expresa cada RF del núcleo funcional de activos, en términos de características, subcaracterísticas y atributos de calidad en concordancia con el estándar de ISO/IEC 25010:2023. También se elabora el listado de los Requisitos de Calidad (RC) de acuerdo a cada RF del dominio estudiado, esto complementa los activos funcionales del dominio.

Salidas:

- Tabla contentiva de los RF, características, subcaracterísticas y atributos de calidad de ISO 25010:2023 [6] asociados al dominio de la familia de sistemas de software.
- Listado de Requisitos de Calidad de acuerdo a cada RF.
- El Modelo de Calidad Funcional (MCF) del dominio de la familia de sistemas de software.

Actividad 5. Descripción de los RNF para Definir el Núcleo de Activos No Funcionales

Entradas:

- ERS de la familia de sistemas de software.

En la actividad se procede a identificar los RNF derivados del dominio estudiado. También se conforma el Núcleo de Activos No Funcionales del dominio.

Salidas:

- Listado de RNF considerando [6], del núcleo de activos no funcionales.

Actividad 6. Descripción y Especificación de la Vista de Calidad No Funcional del Dominio

Entradas:

- Listado de RNF del núcleo de activos no funcionales.

Se especifica la Vista de Calidad del Dominio No Funcional (VCDNF) de acuerdo a cada RNF del núcleo de activos no funcional. Se define cada RNF en términos de características, subcaracterísticas y atributos de acuerdo al estándar ISO/IEC 25010:2023 y se consolida el Modelo de Calidad No Funcional del Dominio de la Familia de Sistemas de Software.

Salidas:

- Tabla de RNF de [6] en términos de características, subcaracterísticas y atributos de calidad del dominio de la familia de sistemas de software incluyendo el análisis.

3. ANÁLISIS DEL DOMINIO EN SPEM

Cada una de las fases de las actividades del Análisis del Dominio IMFS & VCD, que incluyen la creación de artefactos de entradas y salidas que documentan las fases del proceso, se expresan en el modelo SPEM que se presenta en la Figura 3 con sus actividades y especificación detallada de roles, entradas, salidas y estándares utilizados.

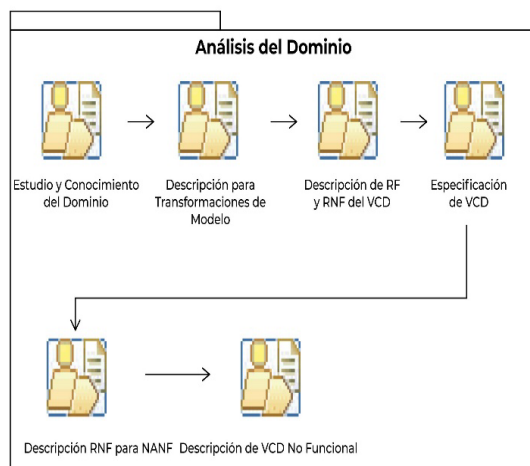


Figura 4: Paquete de la Disciplina “Análisis del Dominio” del Proceso IMFS & VCD

En la Figura 4 se observa la trazabilidad de las actividades que conforman la fase de Análisis del Dominio en términos de SPEM 2.0, siendo importante considerar estos aspectos:

- La caracterización del dominio a través del estudio representado por el modelo de negocio, siguiendo los principios y ciclo de vida iterativo de la gestión de procesos de negocios planteados en [12], con lo que se podrá tener como resultado la información completa del dominio estudiado. Para el desarrollo de esta actividad se necesita la descripción textual y gráfica del problema del dominio para la familia de sistemas de software que se estudia.
- Las transformaciones de modelos que son necesarias desde la perspectiva del estándar MDA, en el marco de la Ingeniería de Modelos, procurando los insumos que son esenciales para cumplir esta tarea: a) El modelado de negocio en BPMN de la actividad previa; b) Los requisitos tecnológicos, c) Los requisitos funcionales y no funcionales que se deriven el estudio del problema del dominio de la familia de sistemas de software.
- Los requisitos funcionales y no funcionales deben ser especificados de acuerdo al dominio de la familia de sistemas de software [15].
- Para la especificación VCD, se deben considerar el ERS [13] completo para el problema de dominio y el modelo de calidad considerado en [6] y finalmente se considera la VCD para los activos no funcionales del dominio.
- Se utilizan para el Análisis del Dominio las disciplinas de:
 - Modelado de negocio.
 - Transformación de modelos
 - Modelado de características de calidad basadas en ISO/IEC 25010:2023.
- Las herramientas, técnicas y estándares utilizadas para aplicar el modelo en cualquier dominio de familias de sistemas de software son:
 - Modelado BPMN.
 - Modelado UML.
 - ISO/IEC 25010:2023.

- IEEE-830-1998.
- IEEE-1471-2000
- Catálogos y estilos de arquitectura de software.
- Experticia del arquitecto de software.

En la Figura 5, se muestra la especificación del paquete “Estudio y Conocimiento del Dominio” de la disciplina Análisis del Dominio de IMFS & VCD [3].

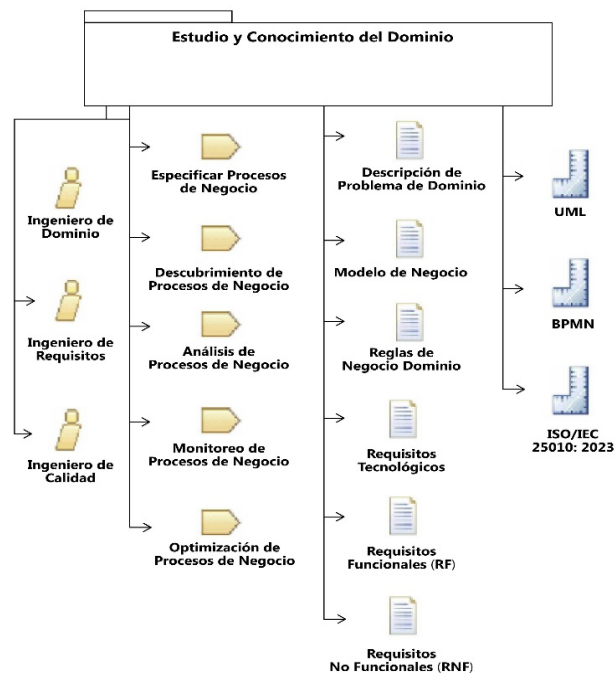


Figura 5: Paquete de Actividad SPEM “Estudio y Conocimiento del Dominio” del Proceso IMFS & VCD

Para cada paquete representado en SPEM existe una descripción de propiedades, ver Tabla II, que permite guiar la aplicación de este proceso para cualquier dominio de estudio de familias de sistemas de software que sea seleccionado.

Tabla II: Descripción de Actividad de IMFS & VCD “Estudio y Conocimiento del Dominio”

Propiedad	Valor
Nombre	Estudio y conocimiento del Dominio.
Rol Responsable	Ingeniero de Dominio, Ingeniero de Requisito, Ingeniero de Calidad.
Descripción	Actividad del proceso IMFS&VCD que trata los aspectos inherentes al problema de dominio y su representación en el modelado de negocio.
Entradas	Descripción de forma textual y gráfica del problema en el contexto de la Familia de Sistemas de Software, para la elaboración de un modelo de negocios específico del dominio.
Salidas	Modelo de Negocio, utilizando la Notación de Modelado de Procesos de Negocio BPMN. Reglas de Negocio, RF y RNF del Dominio.
Técnicas, Herramientas o Estándares	Lineamientos de la Gestión de procesos de Negocios de una organización, BPMN, UML, (SO/IEC 25010:2023.
Disciplina	Construcción del Modelo de Negocio.

4. CONCLUSIONES

La generación de una Arquitectura de Referencia, en el contexto de Familias de Sistemas de Software, dirigida por un proceso que sigue una disciplina rigurosa de la Ingeniería del Dominio

como lo es el Análisis del Dominio, a partir de los enfoques Fábricas de Software, Ingeniería de Modelos y Calidad de Software, es un aporte fundamental a la industrialización de software. En este trabajo se presenta la expresión de este proceso utilizando el lenguaje de modelado para proceso de ingeniería de software SPEM, a través de un modelo de proceso completo *top-down* [14] de descripción arquitectural que está basado en esquemas. El Esquema es uno de los elementos más relevantes en el desarrollo de software bajo el enfoque de Fábricas de Software y en este artículo se ha presentado una parte representativa de esta especificación [3].

Es así como en esta parte representativa, se definen los puntos de vistas que son necesarios para la construcción de un tipo de producto de software o miembro perteneciente a una familia de sistemas de software. Es se logra utilizando transformaciones de modelos y considerando aspectos de calidad que sirven de guía para la obtención de Arquitecturas de Referencia para un dominio industrial de desarrollo de software. Los lineamientos para estas transformaciones de modelos también son aportes de valor para esta línea de investigación en la ingeniería de software.

De manera adicional, se puede señalar que para este modelo de proceso, tanto en el análisis que se muestra en este trabajo como en el diseño, los expertos del dominio y arquitectos de software cuentan con una guía rigurosamente descrita y documentada de las disciplinas, actividades, tareas, artefactos, técnicas y herramientas que permiten que se aplique el proceso de Análisis del Dominio. Contar con esta guía facilita y reduce los tiempos de desarrollo para las aplicaciones de familias de sistemas de software.

Es importante señalar que se requiere de una alta experticia por parte del arquitecto de software para el análisis de los requisitos comunes del núcleo no funcional, de acuerdo con estilos arquitectónicos que garanticen una identificación amplia. Esto es clave para derivar una Arquitectura de Referencia que sea la mejor alternativa de solución con las condiciones del dominio estudiado.

Trabajos Futuros

- Realizar la aplicación del proceso de Análisis del Dominio para arquitecturas de referencias IMFS & VCD modelado con SPEM considerando la actualización del modelo de calidad del producto ISO/IEC 25010:2023.
- Integrar el enfoque ágil para abordar de manera más

efectiva y dinámica la gestión de cambios de requisitos durante la etapa de definición arquitectural.

REFERENCIAS

- [1] IEEE 1471-2000, *IEEE Recommended Practice Architectural Description of Software Intensive Systems*. IEEE, Standards Board, 2000.
- [2] F. Losavio and A. Matteo, *Reference Architecture Design Using Domain Quality View*. Journal of Software Engineering & Methodology, vol. 3, no. 1, 2013.
- [3] A. Arraiz-Goicochea, *Un Modelo de Proceso de Fábricas de Software e Ingeniería de Modelos para la Ingeniería del Dominio Considerando Aspectos de Calidad*. Tesis Doctoral. Facultad de Ciencias, Postgrado de Ciencias de la Computación, UCV, Caracas, Venezuela, Julio 2024.
- [4] OMG, *Unified Modeling Language: Superstructure version 2.0*. <https://www.omg.org/spec/UML/2.0/Superstructure/PDF>. 2004.
- [5] V. Menéndez-Domínguez and M. Castellanos-Bolaños, *SPEM: Software Process Engineering Metamodel*. Revista Latinoamericana de Ingeniería de Software, vol. 3, no. 2, pp. 92-100, ISSN 2314-2642, 2015.
- [6] ISO/IEC 25010, *Systems and Software Engineering. Systems and Software Quality. Requirements and Evaluation (SQuARE)*. Product Quality Model. International Standards, November, 2023.
- [7] C. Ariste, J. Ponisio, L. Nahuel and R. Giandini, *Diseñando Transformaciones de Modelos CIM/PIM: Desde un Enfoque de Negocio Hacia un Enfoque de Sistema*. ASSE 2015, 16° Simposio Argentino de Ingeniería de Software, 2015.
- [8] F. Losavio and A. Matteo, *Reference Architecture Design Using Domain Quality View*. Journal of Software Engineering & Methodology, vol. 3, no. 1, 2013.
- [9] A. Arraiz, F. Losavio y A. Matteo, *Esquema para Fábricas de Software con un Modelo de Vista de Calidad del Dominio*. Tercera Conferencia Nacional de Computación, Informática y Sistemas, CoNCISA, ISBN: 978-980-7683-01-02, Octubre 2015.
- [10] A. Arraiz Goicochea y V. Esteller, *Del Esquema de Fábricas de Software a una Arquitectura de Referencia por Medio de las Transformaciones de Modelos (IMFS & VCD)*, RIMCI, vol. 5, no. 10, pp. 13-31, Julio 2018.
- [11] IEEE-830-1998, *Recommended Practice for Software Requirements Specifications*. IEEE Standards Board, 1998.
- [12] J. Berrocal, J. García, y J. Murillo, *Hacia una Gestión del Proceso Software Dirigida por Procesos de Negocio*. 2007, <http://alarcos.esi.uclm.es/pnis/articulos/pnis-07-Berrocal-GPSDPN.pdf>.
- [13] S. Martínez, L. Lamont, y R. Moreno, *Análisis de la Transformación de Modelo CIM a PIM en el Marco de Desarrollo de la Arquitectura Dirigida por Modelos (MDA)*. Revista Politécnica, vol. 36, no. 1, Septiembre 2015.
- [14] A. Arraiz y Y. Carballo-Barrera, *Aplicación del Proceso IMFS & VCD para Arquitecturas de Referencia de Sistemas Integrados de Salud*. Novena Conferencia Nacional de Computación, Informática y Sistemas CoNCISA. Caracas, Venezuela, Depósito Legal: DC2022001776. ISBN: 978-980-7683-07-4, Noviembre 2022.



Sesión de Artículos Largos

Customizing Software Development Methods: A Process Model Approach

Judith Barrios A¹, Jonás Montilva C¹
ijudith@ula.ve, jmontilva@gmail.com

¹ Department of Computer Science, University of Los Andes, Mérida, Venezuela

Abstract: This article presents an instantiation process model for facilitating the understanding and customization activities of software development methods. The process accepts as an input a method represented by a process model which prescribes the set of software development activities. A method is also defined by a product model that prescribes the set of product parts that can be built by following the process model and, a team model that prescribes the set of roles needed for executing the activities proposed in the process model. The proposal integrates the teaching, practical and consulting experience of the authors, which is essential to understand and handle usual difficulties found during the process of adapting software development methods. The main contribution of our proposal is to provide students of systems and software engineering with a global vision of a method and the savoir-faire implicit in its process model description. Therefore, the proposal simplifies users understanding of method background concepts, and guides them whereas customizing it according to a particular software project context. The proposal is illustrated by an example of the White_Watch method customization to cope with a hypothetical software project situation.

Keywords: Software Methods; Instantiation Process; Method Customization; Teaching Practice

1. INTRODUCTION

Software development methods (SDM) intend to assist project leaders and the whole development team along a development project. A SDM is typically used as a guide to establish main activities as well as technical products that the project team needs to complete to produce a software application. Nevertheless, while following the prescribed guidelines of a method, the project leader -software engineer- has to take instant decisions about the project plan workflow by considering, among other variables, the software product complexity and characteristics, the specific product dynamic requirements, the set of restrictions and needs coming from the current working environment including team size and experience, the developing tools capacities and the essentials of the programming languages, project schedule, and user participation.

In general, a SDM needs some adjustments before being used as a development guide. These adjustments serve to define the project's preliminary schedule and to organize the work that the team must do. Nevertheless, there are still some eventual context factors that may disturb project workflow like technological infrastructure, financial, and human resources complications. A project leader, therefore, needs to frequently adapt the method development process workflow. This problem of tuning a method gets bigger if the project leader/software engineer's background in method understanding and practice is not deep enough. This is the main problem encountered during teaching activities as students begin to understand a method when they must already adapt it to solve a concrete problem. The main motivation of this proposal is, therefore, to assist students to customize a method that they do not completely understand. For that reason, this method instantiation process can be assumed as a teaching practice because it is effective for introducing systems and software engineering students into the context of understanding methods and then adapting them

properly. Nevertheless, the proposition may also help software and systems professionals to understand and customize methods and other methodological guidelines.

The process of fine-tuning or customizing a method is called an Instantiation Process (IP). Generally, an instantiation process is done over the set of general concepts prescribed by a method model or just by a broad methodological textual description – tables or a set of steps or phases. A general method model is prescribed, at a high abstraction level, by a comprehensive set of concepts and their relationships which are involved in the process of developing a particular product or service [1].

In this article, we propose an instantiation process (IP) model for properly guiding the customization of a SDM. It means that the IP accepts a SDM represented by at least a process model, which is one of the three formal method description models. The other two are the product model which prescribes the set of product parts that can be built by following method process guidelines, and the team model which prescribes the set of roles and responsibilities that developers need to take for executing the activities prescribed as process guidelines [1], [2]. This is a generic proposition; thus, it can also be instantiated to couple with methods that only have one or two of the method models mentioned. The instantiation process model is illustrated with the customization of a SDM example that describes how to adapt the method for fitting a particular project situation of a hypothetical example.

The proposal integrates the teaching, practical and consulting experience of the authors as the basis of a good understanding of the problems faced by students and professionals during the adaptation of methods. The feedback gained from teaching system engineering students to customize methods to a specific software project and product factors has been essential. The IP model proposal is completed by ways of working derived from

specific methods adaptations and characterizations completed in many Venezuelan software organizations.

The main contribution of our proposal is to help systems and software engineering students with method understanding and its customization; especially, the concerns related to a software development method adaptation for satisfying specific project needs. In addition, and considering that the SDM already customized may be applied over and over in similar project contexts, students can enhance their ways of working which are perceived as the enhancement of the quality of their development processes as well as that of the software products or services elaborated.

The article is structured as follows: the next section presents the problem of customizing SDM along with the review of some related works. Section 3 presents a summary of method engineering and business process model background concepts. Section 4 describes, at two levels of detail, the instantiation process model proposed; first, a general contextual level and then, the corresponding detailed workflow level for describing the main activities included in the general process model. Section 5 shows how to apply the proposed process by instantiating the White_Watch method in a hypothetical software project situation. Section 6 concludes the paper and gives some practical recommendations to take advantage of the use of the instantiation process model proposed.

2. PROBLEM AND RELATED WORK

2.1 Interpretations of the Problem

The problem of adapting methods is not new, and it is, generally, neglected by software organizations and project leaders since they prefer to solve method adapting problems incidentally [3]. Nevertheless, in preventing to take rush decisions related to tailoring the SDM throughout the execution of a development project, many software organizations define and institutionalize their own development methods to swift project performance and to adequately assure that a high-quality development process shall produce a high-quality software product. But the problem of adapting methods still causes difficulties that disturb project plan work breakdown structure, timelines, costs, and final product or service quality.

The software engineering community early began to work on method selection and its adapting problem. For instance, at the beginning it was identified as a problem of programmers and the knowledge they had of the development cycle along with their capacities to define what to do, when to do it, and how much effort they needed to complete a prescribed task; so, they could organize their work and, consequently, help project leaders to estimate the required work of the whole team. These were the personal software process model (PSP) and TSP [4], [5]. Afterward, the SEI proposal of the CMMI model expected to help organizations and software engineers to understand the whole software development process and, organize team development work by approving a predefined set of business development processes. Accordingly, an organization defines the complete set of software processes that ought to be installed, executed, measured, and monitored to improve organizational performance as well as the corresponding quality of the production process and its products and services [6].

Nevertheless, another customizing problem emerges when a software organization defines or selects, as mandatory, a particular SDM or decides to define and install a set of organizational software processes for elaborating and managing demanded software products. That is, the project leader and the work team must plan and execute each one of the prescribed activities to produce each prescribed technical and management product or document for realizing later, at the end of a project, that some of these activities were not useful because they do not positively contribute to obtaining the final product or service and its documentation, as required by the client. In consequence, there are a lot of time and effort lost, many expenses, and any added value either for the client or for the software organization itself. Many related proposals for software processes assessment and improvement searched to enhance and elevate installed software process performance and quality like SCAMPI, SPICE, ISO 9000 [7], [13].

Some other propositions tackled the problem of adapting and defining the required product development activities by assembling a set of generic ones with or without modifications. This kind of solution was triggered by, among main issues, the software development cycle, the product type, the project management schedules according to pertinent project situational or practical variants, etc. The WATCH suite of methods is a good example of this kind of method proposal [8], [9] as well as the Crystal method family [10]. These propositions preconize a set of guidelines arranged to manage some of the method adapting problems; as a result, a set of method variants is obtained where they can be selected according to the predefined set of project and product features. Other approaches extend generic activities with practical strategies like RUP (Rational Unified Process) and its agile version RUP agile [11], the ASD (Adaptive Software Development) that repeat iteratively an adapted development cycle as a practical strategy to accomplish software product dynamic or uncompleted requirements [12], [13]; and, the SCRUM project management practical approach that has been used as an agile way for organizing the development work for speedy build functional software products [14].

The SEMAT approach extends the range of the approaches reviewed. The OMG Essence standard has been published as the kernel for software engineering methods. It recommends and organizes, in a generic model, a reduced set of essential elements or concepts which are associated with any software production process. This model may well be applied by the development team to define the practical work to be done along with the things to be produced and manipulated in a particular development project [15]. There are many related works where practitioners apply and extend the essentials to cope with development process issues. [16], [17], [18].

Lastly, we add to this variety of methodological research works and practical approaches, the huge number of specialized SDM that have been proposed to help, assist, and guide the development process according to, among other settings:

- the type of product and its complexity,
- the team size, and its experience,
- the tools available and their capacity,
- the time to have a functional version,

- the dynamic change of product requirements, and
- the development technology available.

Some of these propositions recommend their method by including a list of product characteristics, time to have a functional product, team size, and project typical situations where their methods have proved to be effective [19], [20], [21], [22]. The work presented in [23] addresses the methods configuration problem but only for the agile development approach.

None of these methodological propositions and strategies explicitly include a user guide, method use guidelines, or a set of tips to better adjust or customize their proposals. Knowledge level and understanding of a SDM or approach as well as sufficient software development experience seem to play a relevant role in the selection and, subsequently, customization of a SDM to fit the project and contextual factors. The literature review presented in [24] discusses the types of reasons behind software engineers' decisions to select and use software development methods – adoption of methods. It includes discussion about method deviations and method operationalization problems where a solid understanding of SDM background concepts along with the awareness of why and how to implement some of the method activities, seems to be a significant factor to avoid method misconception, and consequently, its misalignment to software project intentions.

We also found that many of these methodological and strategic offers are supported by social network broadcasts and an experienced set of users, programmers, and developers who communicate their practices and recommend effective ways of working like [25] and [26].

2.2 *The Problem from Teaching Experience Perspective*

To complete the description of the problem, we take some examples from our teaching experience in systems and software engineering. For instance, while working on class projects with our students, we notice that central difficulties of method usage were related to the know-how to apply a method. That is, first, students exposed problems to select, and then, how to execute a prescribed guideline/activity or to choose one of them instead of another among the set of the proposed ones. It is not easy for the students to discern if a particular action is necessary (ought to do it), if it is optional or not required considering the type of product/service or the development tool available, or a particular management exigence of the client/teacher; similar difficulty with an earlier configuration of technical products parts that are needed to complete the final software product or service. A concern case is when a student or a project course team decides to follow each one of the prescribed guidelines to produce each one of the prescribed products parts or documents (“because the method is explicit so we have to do it”) for realizing later, at the end of the project, that a lot of the diagrams or executed actions were not necessary because they were redundant or do not contribute at all to the final product/service required.

As observed, the problem of customizing a SDM disrupts software engineering work from many perspectives. It is an individual problem, a team problem, a project leader problem,

and a business process problem. We may conclude that it is not a method adapting problem but a process domain knowledge problem or a deficient practical experience. It may be true but not completely because the process of learning requires a well-founded background to have an understanding and a comprehensive picture of a method before applying it.

That is one of the reasons why this method instantiation process model was shaped: to assist method understanding and comprehension before instantiating it. From the business process perspective, this problem needs to be solved at the institutional level by defining and installing the required and flexible set of software processes by using a specific process improvement approach or standard as mentioned earlier in this section, but this discussion is out of the scope of this proposal.

We are convinced¹ that a generic instantiation process model, like our proposal, may contribute to facilitate the software engineering task of adjusting a SDM to project context situations, product/service characteristics, and team size and experience issues. It means that before starting a development project, the student, the project leader, or the system/software engineer in charge, may analyze and take decisions about the final product or service, its product parts, and documents that must be produced; besides, what would be the workflow that best fit project context situation and team members' knowledge, skills, experiences, and competencies. Method instantiation process model guidelines assist not only to adjust a method to a particular scenario but to understand what is prescribed by the method before adapting it. This premise persists and is independent of the SDM approach, i.e., if it is disciplined, balanced, or agile.

3. BACKGROUND CONCEPTS

Software methods are formally defined as a set of cohesive and complementary models that represent the software final product and its partial components, the software process/activities which need to be executed to produce each part or component of the product/service, and the competencies and understanding that each member of the software team must attest to, adequately, execute each prescribed activity, and assure that the product part elaborated has the expected quality. These models are the product model, the process model, and the team model, respectively [1], [2].

The process of fine-tuning or customizing a method is called an instantiation process (IP). Generally, this kind of process is done over a set of generic concepts prescribed by a method model. This method model is general and is placed at a high abstraction level. It means that a software development model prescribes the complete set of concepts and their relationships which are involved in the method development process of a software product/service. Some of them may as well include the roles that the members of a software team have to play throughout a development project.

The process of instantiating a method implies the selection, extension, reduction, or modification of any of the concepts included in the general model. In the case of a SDM represented by a process model, a product model, and a team model, the instantiation process ought to be done coherently and

¹ After applying several times some others associated learning strategies

consistently over each one of the generic models so the relationships and dependencies between concepts and models may be correctly kept. After an instantiation process, the customized method model has a step lower abstraction level than the general one.

Figure 1 shows the links between the generic method model and the customized model obtained after an instantiation process.

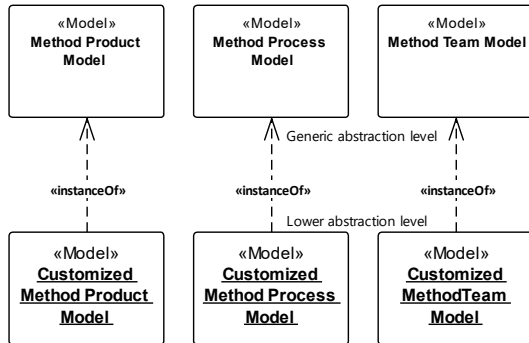


Figure 1: Links between Generic and Instantiated Method Models

A method product model represents the set of product parts and the relationships among them which can be elaborated by applying a particular method process guideline. In the case of a SDM, a method model must include the set of partial technical and management product parts as well as the deliverable ones. Figures 2, 3, and 4 present at a very high abstraction level a set of generic method concepts that may be instantiated to build a particular SDM product model. These concepts can be instantiated to fit specific method requirements [8], [9].

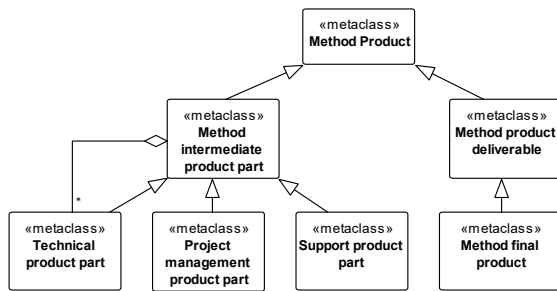


Figure 2: Generic Product Model Concepts. Adapted from [8]

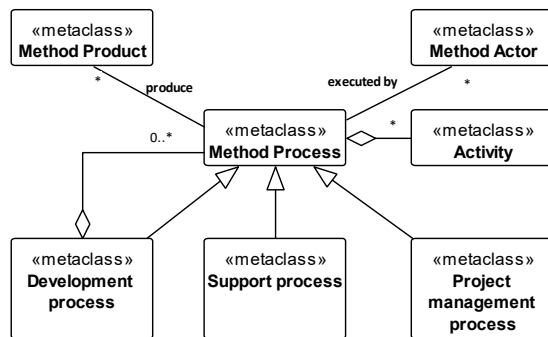


Figure 3: Generic Process Model Concepts. Adapted from [8]

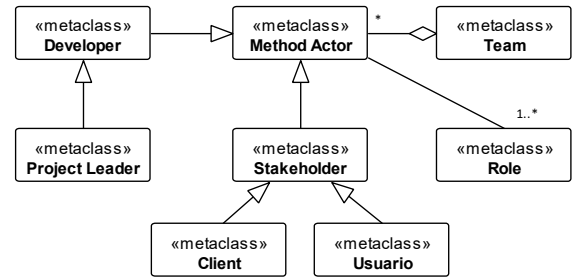


Figure 4: Generic Team Model Concepts. Adapted from [8]

As model concepts are represented at a high abstraction level, they could be chosen and extended for being part of a particular SDM. These meta-models were used to elaborate the Blue, Yellow, and White variants of the WATCH method suite [9].

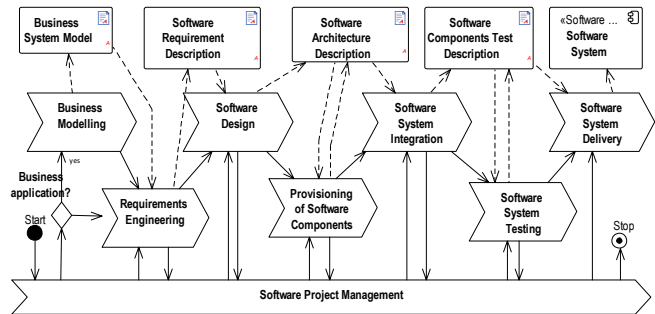


Figure 5: Workflow of the White_Watch Method [27]

Figure 5 presents an example of the results of the instantiation process of the workflow of the White_Watch method [27]. This version of the WATCH suite was defined for assisting software engineering students in their course projects. Notice that product and process generic concepts (Figures 2 and 3) were extended to have a more specific description of the technical products and their corresponding building processes. In this case, a software product is built as an assembled set of reused software components. The intermediate technical products of Figure 2 were extended with a business system model, documents for expressing software requirements, architecture, and testing descriptions. These technical products are represented by using explicit software engineering techniques and UML diagrams where correspond.

3.1 Process Model

The IP model proposed is graphically represented by a business process model using the UML Business notation [28]. Accordingly, a business process explicitly has a process goal to aim, is executed and supervised by actors (and their roles), has a set of inputs that may be transformed into outputs, generates other outputs, and is regulated by some precise procedures, standards, and rules, and is supported by some business resources (technology, money, documents, etc.). A business process may be complex or simple so it can be decomposed and detailed (into activities and tasks) according to modeling description requirements. A graphical representation like business process models is easy to understand and follow and provides students and engineers with a complete perspective of

what they must do for customizing a software development method [29] and [30].

For that reason, the IP model is represented by a general process description diagram and a process decomposition diagram that shows the set of four sub-processes of the general process. Each one of the sub-processes is detailed by using a UML activity diagram.

4. THE INSTANTIATION PROCESS MODEL

As we explained earlier, an instantiation process implies the selection, extension, reduction, or modification of any method element included in the general method model. This instantiation aims to generate a particular version of the method. Besides, in the case of a SDM represented by a process model, a product model, and a team model, the instantiation process ought to be done coherently and consistently over each one of the general models so the relationships and dependencies between concepts and models can be respected.

According to the research works [8] and [9], method guidelines suggest starting the instantiation process by first customizing the method product model determining what is going to be produced by the adapted method; then, the selection of the process model elements that describe what must be done to produce the product model elements already selected. The instantiation process ends by defining the actors and their roles that are required for executing those processes/activities expressed by the process model already instantiated. Afterward, a validation process is necessary to assure that the resulting method models are coherent among them. According to these process model guidelines, the student/project leader has by now shaped a set of method models to a particular software project scenario.

Figure 6 shows the Instantiation Process description by using a general-level UML Business diagram [28].

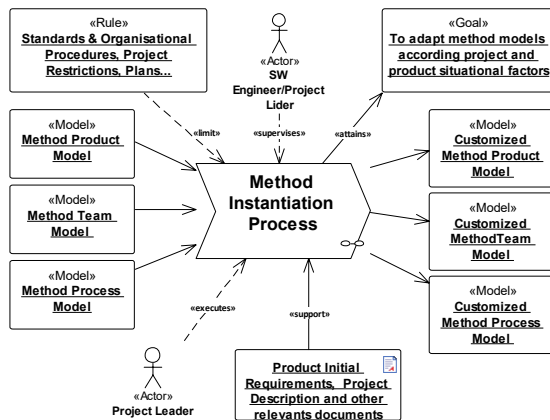


Figure 6: General Description Diagram for the Instantiation Process (IP)

As described in the previous diagram, the IP is modeled as a business process whose main goal is “to adapt method models according to project and product situational factors”. The IP process accepts general Method Models (Product, Process, and Team) and after processes guidelines, it produces the set of corresponding customized method models. The process is under

the responsibility of the project leader (actor) and it may be supervised by a software engineer. There are some rules, standards, and restrictions related to method application, organizational or business domain, and other method features that must be considered while customizing method models. Project documents and other items like initial product requirements, organizational context, and some relevant technological parameters support decisions related to the selection of model elements. Similarly, if there are any required addition or modification to complete general method model customization.

Considering that the IP is a complex process it has been decomposed into four sub-processes as represented in Figure 7.

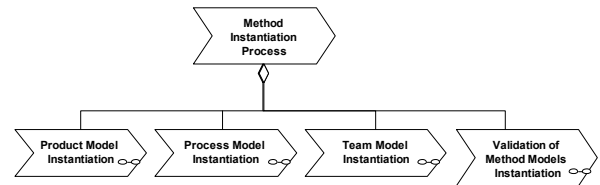


Figure 7: Sub-processes that Compose the IP

To precisely prescribe IP model guidelines, the detailed workflow for each one of the four sub-processes depicted in Figure 7 is presented in Figures 8, 9, 10, and 11, respectively.

According to Figure 8, for customizing a method product model, it is necessary to analyze the initial software product requirements document as well as consider project contextual and technological aspects that may influence the selection and characterization of a set of product parts to be produced as represented in the workflow of the figure. It may be necessary to add or modify by extension or reduction one or more product parts of the customized product model. The IP of the method product model ends after a well-structured validation of the whole set of method product elements.

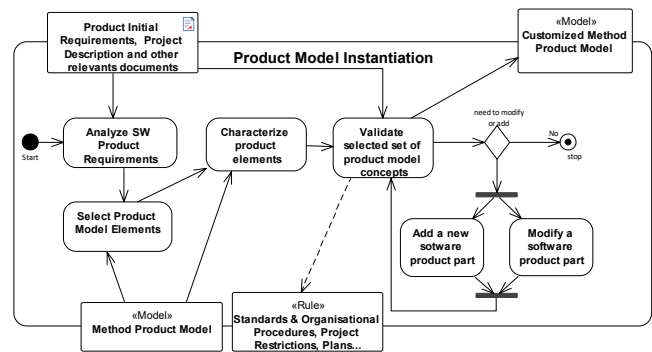


Figure 8: Product Model Instantiation

Once the method product model is customized, the IP continues with the method process model to select the required set of processes/activities that are required for building each one of the product model elements included in the customized product model. If there are some new product parts or some of them have been modified, the corresponding set of method process model elements should be defined or redefined as appropriate. It is possible that the general method process model does not have all the prescribed processes, i.e., maybe it just prescribes

the development processes but not the support or the project management processes. Therefore, the IP is done only on the available process model elements as represented in the activity diagram of Figure 9. It is important to have in mind that any modification to current method models must be properly described by using the same formality and/or notation that in the SDM general model.

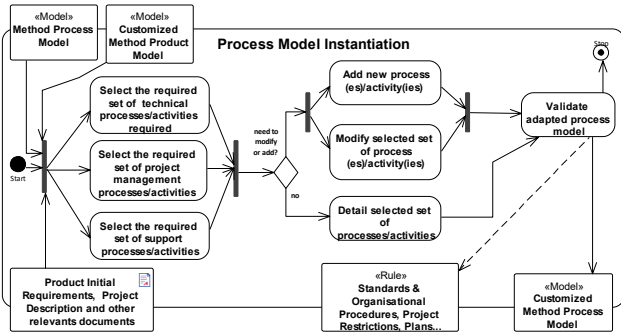


Figure 9: Process Model Instantiation

If a method team model exists in the general SDM models, it is instantiated by using as inputs the already customized method process model. This process consists of identifying and characterizing actors, roles, and responsibilities required for the execution of processes/activities included in the method process customized model as it is represented in Figure 10.

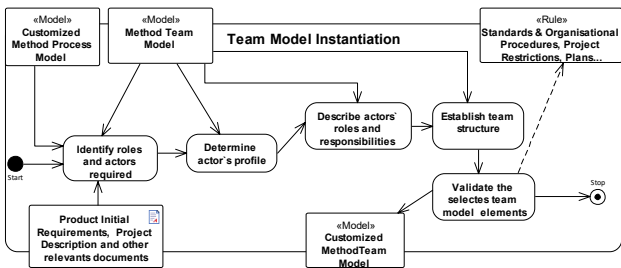


Figure 10: Team Model Instantiation

Finally, a global validation process is done to assure that the set of customized models is coherent and consistent with what is expressed in the SDM general models, and with the project and technological factors delimited by the official project documents, procedures, rules, and restrictions (see Figure 11). For example, in a SDM expressed by the three method models, the process model must use, produce, or complete a product part included in the product model. In the same way, each element included in the process model must be described or assigned as a responsibility of an actor's role from the team model.

As we mentioned, the proposal is a generic process model, thus it is adjusted to different kinds of method descriptions and formalities. As mentioned before, the proposed IP requires, as input, at least a process model. Most methodological descriptions are represented textually through tables or as a list of steps or phases. This type of representation expresses activities that are performed to produce a software product (partial or complete). Therefore, the product model is implicitly included in such activities; it is the task of the project leader to extract the products involved, and to determine what part of the product

they may represent and whether or not they form part of the deliverable product. The general IP workflow is showed in Figure 12.

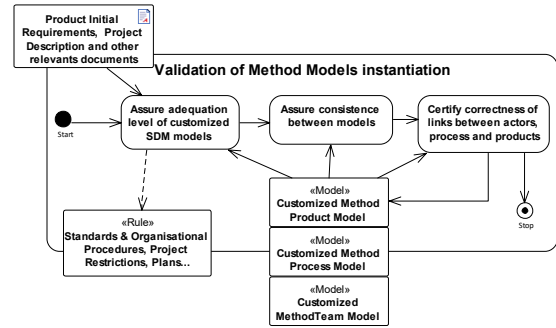


Figure 11: Validation of the Customized Set of Method Models

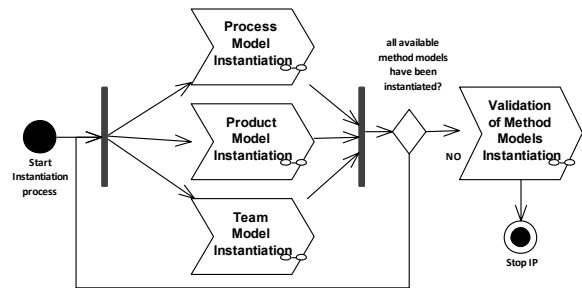


Figure 12: Instantiation Process General Workflow

5. APPLYING THE INSTANTIATION PROCESS MODEL

To illustrate the benefits of our proposal, we present the customization of the White_Watch method [27] for a hypothetical course project where the students work in a team of two software engineers.

5.1 Hypothetical Project Description

The first part of a software engineering course project consists in detailing the set of business processes involved in an online booking system for a small theater. These processes are part of the organizational or business system model of the small theater; thus, it is necessary to define what are the activities, actors, events, resources and objects involved in the process of booking seats in the plays offered.

5.2 IP Workflow for Adapting the White_Watch Method

The White_Watch method workflow showed in Figure 5 is detailed through a descriptive table that organizes those processes into steps. Each step has a set of prescribed activities along with the set of techniques or notations suggested to elaborate the involved products. In view of that, students should adapt the method from the table description which has four columns: method steps, activities, notations/techniques and products. Table I shows the Business System modeling step of the White_Watch method.

Initially, it is important to state that, according to Figure 6, both students, alternatively, ought to play the role of project leader during the personalization of the process prescribed in the IP model.

Table I: Excerpt of White_Watch Method Model [27]

Steps	Activities	Notations/techniques	Products
Business System (BS) Modeling	-Modeling BS value chain (if needed)	-Interview with BS and domain experts	-Value Chain
		-Direct observation of BS context	-Hierarchy of processes
	-Modeling fundamental processes	-Review of technical documentation	-Descriptions of Processes
	-Modeling Support processes	-UML Business value chain diagram	-Diagrams of activities
	-Modeling process' activities	-UML Business process description diagram	
		-UML activity diagram	

The relationship between product and process models is explicit and direct. Students have to select and instantiate those products, steps, and activities considering the above-mentioned course project requirements; and then they can select the suggested technique or notation to produce them. IP guidelines suggest to start the instantiation process by the product model, then the process model, and finally, if required, the team model.

Table II: White_Watch Method Model Customized for the Example

Steps	Activities	Notations/techniques	Products
Business System (BS) Modeling	-Modeling fundamental processes	-Direct observation of BS context	-Descriptions of Processes
		-UML Business process description diagram	-Diagrams of activities
	-Modeling process' activities	-UML activity diagram	

The outcome of the IP execution should be:

- Product model required (from Table I column “products”): one or more business processes descriptions and the corresponding set of activity diagrams.
- Process Model instantiation (Table I column “activities”): *modeling fundamental processes* (for the booking process) and *modeling process activities* (for detailing the booking process). These activities should be performed by applying techniques and notations selected from the " Notations/ Techniques" column. In that case, after *direct observation of the business context*, it is necessary to represent perceived processes and activities by using UML business process description diagrams and UML activity diagrams.
- The adaptation of the Team Model seeks to define the roles that each of the two students has to play to execute the process model already instantiated. The actor roles prescribed by the method are “project leader”, “analyst”, “designer”, “programmer” and “tester”. In this case and as part of the teaching strategy, both students must play the role of *analyst* at some point in the modeling process.

Table II presents the customization of the White_Watch method after applying the IP to the example introduced above.

6. CONCLUSIONS

Throughout this article we presented an IP model as a teaching strategy to help students while personalizing a method they do

not fully understand. The proposed model provides students with a more complete view of the components of the method and their dependency relationships, which are essential to adapt a method to the situation of a particular project. In fact, it allows to better apprehend what is the purpose of a method and how it is structured, the kind of product that can be built through the execution of a pertinent set of processes and activities, and to detect what are the competencies that developers need to have to properly implement prescribed activities.

As explained, the problem of adapting methods has many relative solutions; some of them are oriented to the implementation of method variants that partially solve the problem of choosing a suitable SDM according to some predefined factors. Others, characterize the methods according to their development approach to then suggest how to select a method from the collection. Our proposal is generic and is based on method engineering concepts so it can be adjusted to different types of method descriptions and formalities, independently of any approach, paradigm and type of method. For instance, it equally works for disciplined, balanced or agile development methods.

This proposal has proven to be an effective teaching practice for introducing system and software engineering students into the conceptual context of the development methods and their adaptation. The IP model may also assist systems professionals to understand and customize other methods and methodological guidelines than SDM.

REFERENCES

- [1] S. Brinkkemper, “*Method Engineering: Engineering of Information Systems Development Methods and Tools*”. Information and Software Technology, 38, 275—280, 1996.
- [2] J.J. Odell, “*A Primer to Method Engineering*”. INFOSYS: The electronic newsletter for information systems. 3 (19), 1996.
- [3] D. Rivero, J. Montilva, J. Barrios and M. Murúa, “*Un Análisis del Desarrollo de Software en Empresas Venezolanas*”. Seventh LACCEI Latin American and Caribbean Conference for Engineering and Technology- LACCEI'2009. San Cristobal, Venezuela, pp. WE1: 1-10, 2009.
- [4] W. Humphrey, “*Introduction to the Personal Software Process*”, Addison-Wesley, 1997
- [5] W. Humphrey, “*Introduction to the Team Software Process*”, Addison-Wesley, 2000
- [6] SEI “*CMMI for Development, Version 1.3*”. Software Engineering Institute. Carnegie Mellon University, USA, Technical Report No.CMU/SEI-2010-TR-033, 2010.
- [7] M. de la Villa, M. Ruiz and I. Ramos, “*Un Estudio Crítico Comparativo de ISO 9001, CMMI e ISO 15504*”. CISTI 2006 ISBN: 978-989-20-0271-2 Volume II, pp. 551-551551. 2006. <https://www.researchgate.net/publication/235661307>
- [8] J. Barrios, J. Montilva and D. Rivero, “*The WATCH Method Suite in Practice: Two Complementary Perspectives of Use*”. Conference CLEI 2011. Quito Ecuador. Octubre, 2011.
- [9] J. Barrios y J. Montilva, “*Watch: A Suite of Methods for Facilitating Software Development Process Adaptability*,” in Software Engineering: Methods, Modeling, and Teaching. Sello Editorial Universidad de Medellín. Colombia, 2011.
- [10] A. Cockburn, “*Crystal Clear*”, Addison-Wesley, 2005.
- [11] S. Ambler, “*The Agile Unified Process (AUP)*”, 2006. <http://www.ambysoft.com/unifiedprocess/agileUP.html>
- [12] J. Highsmith, “*Adaptive Software Development: An Evolutionary Approach to Managing Complex Systems*”, Dorset House Publishing, 2000.

- [13] R. Pressman, “*Software Engineering a Practitioner’s Approach*”. Seventh Ed. by The McGraw-Hill Companies, Inc, 2010.
- [14] K. Schwaber and J. Sutherland, “*The Scrum Guide - The Definitive Guide to Scrum: The Rules of the Game*”. Scrum.org. October, 2017. <https://www.scrumalliance.org/learn-about-scrum/the-scrum-guide>
- [15] I. Jacobson, Ng Pan-Wei, P. McMahon, I. Spence, and S. Lidman, “*The Essence of Software Engineering: The SEMAT Kernel*”. Communications of the ACM, 2012, vol. 55, no. 12, December, pp. 42-49.
- [16] *Essence Kernel and Language for Software Engineering Methods Version 1.0*, Nov. 2014. <http://www.omg.org/spec/Essence/1.0/>
- [17] V. Savić, E. Varga, “*Extending the SEMAT Kernel with the TDD practice - IET Software* - Wiley Online Library, 2018 <https://doi.org/10.1049/iet-sen.2016.0305>
- [18] J. S. Park, “*Essence-Based, Goal-Driven Adaptive Software Engineering*,” IEEE/ACM 4th SEMAT Workshop on a General Theory of Software Engineering, 2015, pp. 33-38, DOI: 10.1109/GTSE.2015.12
- [19] J. Montilva and J. Barrios, “*Ingeniería del Software: Un enfoque basado en procesos*”. Sello Editorial del Vicerrectorado Académico, Universidad de los Andes, Venezuela. August, 2021. Published by Amazon Books. Available: <https://www.amazon.com/-/es/Jon%C3%A1s-Montilva-C/dp/9801120290>
- [20] J. Barrios and J. Montilva, “*A Balanced and Adaptable Method for Software Development in very Small Enterprises: The Blue Watch Variant*”. In Carlos M. Zapata, et al (eds), *Software Engineering: Methods, Modeling and Teaching*: 39-54. Medellín: Sello Editorial. 2011.
- [21] J. Barrios and J. Montilva, “*Gestión Ágil de Proyectos a través de la Integración de Blue-WATCH y SCRUM*”. X Ibero-American Conference on Software Engineering and Knowledge Engineering – IIISIC 2013. University of Medellín. Medellín, Colombia. November, 2013.
- [22] M. A. Khan, A. Parveen and M. Sadiq, “*A Method for the Selection of Software Development Life Cycle Models Using Analytic Hierarchy Process*,” International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT), 2014, pp. 534-540, doi: 10.1109/ICICT.2014.6781338.
- [23] D. Gupta and R. Dwivedi, “*Configurable Method Model of Agile Methods for Creating Project-Specific Methods*”. Int'l Conf. Software Eng. Research and Practice, SERP'16. ISBN: 1-60132-446-4, CSREA Press. 2016
- [24] T. Havstorm, F. Karlsson. “*Software Developers Reasoning Behind Adoption and Use of Software Development Methods – A Systematic Literature Review*”, in International Journal of Information Systems and Project Management. Vol 11, No. 2, Article 4, 2023. <https://aisel.aisnet.org/ijispm/vol11/iss2/4>
- [25] *Toptal Engineering Blog*. <https://www.toptal.com/developers/blog>
- [26] SoVisionIT. *Software and APP Development*. <https://www.sovisionit.com/software-app-development/interesting-stuff/blog/>
- [27] J. Barrios and J. Montilva, “*Método W Watch*”. Revised version 2020. Internal Report, Group GIDyC. School of Systems Engineering. University of Los Andes. Mérida, Venezuela. 2020. Partially published in J. Montilva & J. Barrios [21].
- [28] H.E. Eriksson and M. Penker, “*Business Modeling with UML: Business Patterns at Work*”. John Wiley & Sons, 2000.
- [29] K. Figl, “*Comprehension of Procedural Visual Business Process Models: A Literature Review*”. Business and Information Systems Engineering Vol 59(1), February 2017, 41–67.
- [30] E. Kornysheva and J. Barrios, “*Visual Representation of the TOGAF Requirements Management Process*”. ER Workshops 2225 October 2018. LNCS, volume 11158, pages 239–248.

Estrategias para Impulsar el Ecosistema Emprendedor Tecnológico Venezolano: Análisis DOFA Ampliado

José Fernández¹, Yusneyi Carballo-Barrera¹
jose.fernandez.v@ucv.ve, yusneyicarballo@gmail.com

¹ Escuela de Computación, Universidad Central de Venezuela, Caracas, Venezuela

Resumen: El artículo analiza los desafíos y oportunidades del ecosistema emprendedor tecnológico en Venezuela, a partir de la revisión bibliográfica y utilizando un modelo sistémico en donde se evalúan tres ejes principales: capital humano, espacio de oportunidades, y estructuras de soporte para el emprendimiento. Se desarrolla un análisis de matriz DOFA a fin de identificar factores clave asociados a debilidades, oportunidades, fortalezas y amenazas del ecosistema emprendedor en el país. Entre las debilidades destacan la falta de innovación y un marco regulatorio inadecuado, mientras que las oportunidades incluyen iniciativas locales exitosas y el interés en la transformación digital. Las fortalezas se relacionan con la formación técnica de las emprendedoras y de los emprendedores, además de la existencia de una plataforma de ciencia, tecnología e innovación. Como amenazas para el ecosistema emprendedor destaca la percepción de inestabilidad en el país, la desconfianza, la falta de financiamiento y el desestímulo al consumo. A partir del análisis cruzado de los elementos de la matriz DOFA, se proponen algunas estrategias para fortalecer el ecosistema de base científica y tecnológica, además de impulsar la creación de empresas de alto impacto en Venezuela. A pesar de los desafíos, se identifican estrategias cuya implementación consideramos como un factor clave para el crecimiento económico y de la innovación en el país, en donde la cooperación intersectorial e interinstitucional es fundamental.

Palabras clave: Ecosistema emprendedor; Tecnología; Emprendimientos de Base Científica y Tecnológica en Venezuela; Emprendimientos Dinámicos; Dimensiones del Ecosistema Emprendedor; Matriz DOFA ampliada.

1. INTRODUCCIÓN

El ecosistema emprendedor tecnológico, entendido éste como un conjunto de interacciones entre emprendedores, instituciones y recursos que fomentan la innovación y el crecimiento organizacional, enfrenta en Venezuela múltiples desafíos y oportunidades en un contexto de crisis multifactorial, donde el 96 % de los emprendimientos son del tipo de subsistencia o necesidad [1].

El modelo de visión sistémica del ecosistema emprendedor, desarrollada por Kantis y Federico [2], contiene tres ejes (Capital humano, Espacio de oportunidades, Estructuras de soporte) en los que se distribuyen diez (10) dimensiones que se pueden evaluar por separado para determinar qué tan favorable o desfavorable es la situación para los emprendimientos tecnológicos, como una categoría que forma parte de los emprendimientos dinámicos.

Este artículo resume el resultado del análisis estratégico del ecosistema emprendedor tecnológico venezolano, considerando tanto factores internos y como externos, para luego formular un conjunto de estrategias para impulsarlo. Estas emergen de los elementos identificados en una matriz DOFA ampliada (matriz de Debilidades, Oportunidades, Fortalezas y Amenazas). Se espera que el conjunto de estrategias propuestas contribuyan a conformar un marco de referencia para fomentar el desarrollo de emprendimientos de base científica y tecnológica (EBCT) en Venezuela, los cuales puedan a su vez impulsar la creación de empresas de alto impacto y acelerar el surgimiento de “unicornios”, es decir, las empresas en el sector tecnológico y de la innovación (*startups*) con valoraciones superiores a los mil millones de dólares.

2. MARCO TEÓRICO

2.1 *Emprendimientos de Base Científica y Tecnológica*

De acuerdo a un informe realizado por Kantis y Angelelli [3] para el Banco Interamericano de Desarrollo se definen a los Emprendimientos de Base Científica y Tecnológica (EBCT) como organizaciones cuyo desarrollo está basado mayoritariamente en intangibles como el conocimiento proveniente de actividades de investigación y desarrollo (I+D). Es un tipo de organización que impulsa el crecimiento económico, la innovación y el bienestar social, de ahí el interés de los países en el desarrollo de este tipo de emprendimientos. Bajo el enfoque de los autores antes mencionados, el papel de las instituciones de educación superior o universitaria, los centros de investigación y las actividades privadas de I+D son fundamentales para el desarrollo del EBCT.

La clasificación de los emprendimientos en cuatro tipos principales (De supervivencia, estilo de vida, crecimiento gestionado y de crecimiento agresivo) desarrollada por Morris *et al.* [4] da pie, en el marco de esta investigación, para la discusión acerca de las condiciones necesarias para el crecimiento de uno u otro tipo. Otro enfoque cuya exploración también es importante en este trabajo es el “emprendimiento dinámico”, el cual ha sido definido por Hidalgo, Kamiya y Reyes como “*iniciativas empresariales impulsadas por equipos emprendedores que identifican una oportunidad de negocio con alto potencial de crecimiento, a escala regional o global*” [5, p. 13]. Es un enfoque importante al cual los autores han dedicado varias publicaciones promovidas por la Corporación Andina de Fomento. Siguiendo con las clasificaciones, otros autores destacan que los emprendimientos de base científica y

tecnológica por su comportamiento y características generales, entran en la categoría de emprendimientos dinámicos [2][3].

2.2 Ecosistema Emprendedor

El Ecosistema Emprendedor, según Kantis y Federico [2] se define como un conjunto dinámico de actores, instituciones y procesos que interactúan para fomentar la creación y el crecimiento de nuevas empresas. Este ecosistema se caracteriza por su capacidad de adaptación y evolución, donde el Estado juega un papel crucial en la creación de condiciones propicias para el emprendimiento, como la provisión de capital inicial y el apoyo a incubadoras. A medida que el ecosistema madura, el liderazgo se transfiere al sector privado, aunque el Estado mantiene un rol de apoyo clave.

Los autores describen en su modelo cómo las dinámicas complejas y sistémicas que ocurren en la interacción de los diferentes actores, aunadas a las fuerzas transformadoras, crean una atmósfera de condiciones favorables para el crecimiento, el desarrollo y la evolución. Las dinámicas y las fuerzas transformadoras crean un emprendimiento dinámico que incluye a empresas de alto impacto con crecimiento sostenido y exponencial (empresas gacelas), pero también, a iniciativas de emprendimiento que logran sobrevivir los primeros años, para luego convertirse en poco tiempo en pymes o empresas competitivas y con proyección de crecimiento [6].

Los elementos característicos del ecosistema emprendedor pueden ser mejor entendidos gracias a la metáfora de Daniel Isenberg [7], la cual compara las características de un ecosistema natural con un ecosistema emprendedor. La Tabla I muestra esta comparación, destacando las diferencias y las similitudes entre ambos ecosistemas, enfatizando cómo cada uno opera dentro de sus respectivos contextos y objetivos.

Tabla I: Comparativa en el Ecosistema Natural y el Ecosistema Emprendedor Según la Metáfora de Isenberg

Características	Ecosistema Natural	Ecosistema Emprendedor
Componentes	Organismos, clima, suelo, agua.	Empresas, emprendedores, instituciones, capital.
Interacciones	Relaciones simbióticas, depredación, competencia.	Colaboración, competencia, redes de apoyo.
Dinamismo	Adaptación a cambios ambientales.	Adaptación a cambios del mercado y tecnología.
Objetivo	Supervivencia y reproducción de especies.	Crecimiento y sostenibilidad de empresas.
Recursos	Recursos naturales renovables y no renovables.	Capital humano, financiero, tecnológico.

2.3 Comprensión del Ecosistema Emprendedor

Para la comprensión del ecosistema emprendedor es necesario el análisis de sus condiciones ecosistémicas, las cuales se pueden organizar en tres (3) ejes conceptuales y en diez (10) dimensiones. Los tres ejes son [8]:

1. *Capital humano emprendedor y sus ámbitos formativos:* Se refiere a la educación y habilidades de los emprendedores.

2. *Espacio de oportunidades y sus determinantes:* Incluye los factores que crean y limitan oportunidades para el emprendimiento.
3. *Estructuras de apoyo o soporte:* Destaca la importancia de las relaciones y colaboraciones dentro del ecosistema.

El modelo se sirve, al mismo tiempo, de 10 dimensiones que proporciona un marco integral para entender cómo los diferentes elementos interactúan y afectan la capacidad de un ecosistema para fomentar el emprendimiento y la innovación (ver Figura 1). Las dimensiones del modelo o condiciones marco según el modelo de Kantis, Federico e Ibarra son [6][8]:

1. *Capital Humano Emprendedor:* Son las personas con la vocación y la motivación para emprender. Constituyen una masa crítica necesaria para que los emprendimientos logren recorrer su sendero de crecimiento.
2. *Cultura:* Se refiere a los valores, creencias y actitudes que predominan en una sociedad y que pueden fomentar o inhibir el emprendimiento.
3. *Educación:* La educación tiene un papel fundamental en la formación de capital humano emprendedor, en especial, la educación universitaria. Las instituciones educativas tienen la facultad de poder fomentar habilidades emprendedoras a través de programas específicos de estudio, promoción y acompañamiento de proyectos de emprendimiento universitario, con una oferta educativa pertinente y de calidad.
4. *Condiciones Sociales:* Analiza el contexto social que afecta el emprendimiento, como la cohesión social y la inclusión.
5. *Estructura Empresarial:* Examina la composición y características de las empresas en el ecosistema, incluyendo su tamaño y sector.
6. *Condiciones de la demanda:* Identificar la existencia de oportunidades de negocios atractivas que estimulan la dinámica emprendedora y empresarial. Estas oportunidades incluyen la presencia de clientes sofisticados, la apertura de mercados externos, y la existencia de sectores dinámicos y en crecimiento.
7. *Plataforma de Ciencia y Tecnología para la Innovación (CTI):* La plataforma CTI incluye las instituciones y recursos que apoyan la investigación y el desarrollo, vitales para la innovación.
8. *Capital Social:* Se refiere a las redes de relaciones y conexiones que facilitan la colaboración y el intercambio de recursos entre los actores del ecosistema.
9. *Financiamiento:* El acceso a diferentes fuentes de financiamiento, desde capital semilla hasta capital de riesgo, es crítico para que los emprendimientos puedan crecer y escalar.
10. *Políticas y regulaciones:* Incluye políticas de fomento al emprendimiento, incentivos fiscales, protección de la propiedad intelectual, entre otros, donde el marco regulatorio y de políticas públicas favorables a la promoción, incentivo y protección del emprendimiento es fundamental.

Factores que pueden promover o inhibir la creación y crecimiento de las empresas

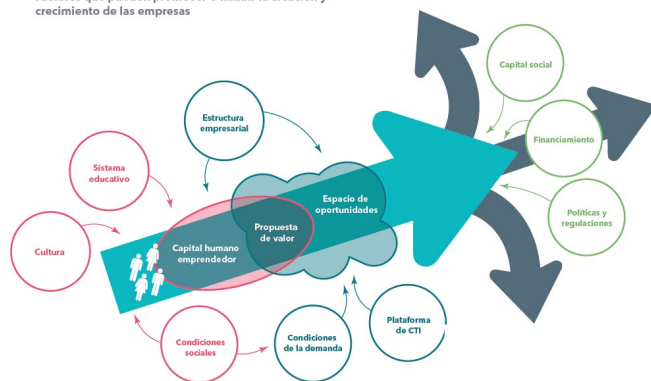


Figura 1: Dimensiones del Ecosistema Emprendedor Según el Modelo de Kantis, Federico e Ibarra

Entre los tres ejes conceptuales y las diez dimensiones se generan varias dinámicas importantes que establecen interacciones sistémicas y moldean la trayectoria particular dentro de cada ecosistema, algunas de ellas son:

- **Dinámica Emprendedora:** Evalúa la actividad emprendedora en el ecosistema, incluyendo la creación de nuevas empresas y la innovación.
- **Dinámica Empresarial:** Se centra en el crecimiento y desarrollo de las empresas existentes, así como en su capacidad de adaptación.
- **Dinámica Institucional:** Analiza el papel de las instituciones en la regulación y apoyo al emprendimiento.
- **Dinámica de Inversión y Financiamiento:** Examina el flujo de capital hacia nuevas empresas y la disponibilidad de financiamiento.
- **Dinámica de Gobierno, Políticas y Regulaciones:** Refiere a las políticas públicas y la gobernanza que afectan el entorno emprendedor, incluyendo la creación de un marco regulatorio favorable.

Las dinámicas son impulsoras clave del sistema emprendedor y se retroalimentan con las otras dinámicas del ecosistema. Esto permite señalar que hay una evolución del ecosistema como un proceso de autorreforzamiento, donde las interacciones o dinámicas entre ejes y dimensiones co-evolucionan a lo largo del tiempo.

3. MARCO METODOLÓGICO

Para formular el conjunto de estrategias con el objeto de fomentar el desarrollo de emprendimientos de base científica y tecnológica e impulsar la creación de empresas de alto impacto en Venezuela se utilizó como herramienta una matriz DOFA ampliada.

3.1 Recopilación de Información

Se realizó una revisión exhaustiva de fuentes de información, generadas entre los años 2010 y 2024, centradas en analizar la situación actual del emprendimiento y la innovación tecnológica en Venezuela, identificando tanto desafíos como oportunidades para su desarrollo. Las fuentes consultadas y analizadas incluyen:

- Artículos científicos y académicos sobre emprendimiento

tecnológico, ecosistemas emprendedores y la situación en Venezuela.

- Informes gubernamentales, de organismos internacionales y de organizaciones especializadas en emprendimiento.
- Datos estadísticos de estudios que analizan el estado del emprendimiento en el país.
- Datos estadísticos de estudios que proporcionan indicadores e información representativa de la situación económica, social y de salud en Venezuela.

3.2 Análisis de Variables Ecosistémicas

Basándose en el modelo teórico de Kantis y Federico [2] se analizaron las variables clave que conforman el ecosistema de emprendimiento dinámico: a) Las condiciones marco del modelo, b) Las dinámicas o impulsores del desarrollo de los ecosistemas emprendedores, y c) La evolución del ecosistema emprendedor como un proceso de autorreforzamiento.

Para el análisis desarrollado en esta investigación las variables ecosistémicas clave se agruparon en tres ejes:

a) Capital Humano:

- Condiciones sociales.
- Capital humano emprendedor.
- Cultura.
- Sistema educativo

b) Espacio de Oportunidades:

- Condiciones de demanda.
- Plataforma de Ciencia, Tecnología e Innovación (PCTI).
- Estructura empresarial.

c) Estructuras de Soporte:

- Capital social.
- Políticas y regulaciones.
- Financiamiento.

3.3 Matriz DOFA y Matriz DOFA Ampliada

La matriz DOFA es una herramienta de análisis estratégico que permite identificar a cuatro factores: 1. Debilidades, 2. Oportunidades, 3. Fortalezas y, 4. Amenazas. También conocida como FODA o SWOT (*Strengths, Weaknesses, Opportunities, Threats*), su utilidad radica en proporcionar una visión integral de la situación de la organización, facilitando la toma de decisiones estratégicas y el establecimiento de objetivos realistas y alcanzables.

La matriz DOFA es una herramienta alternativa útil para diagnosticar y determinar estrategias de intervención en una organización. En el marco de esta investigación, se aplicó para analizar el contexto de factores internos y de factores externos de los emprendimientos de base científica y tecnológica. De este análisis deriva la identificación puntos de intervención y de mejora para el ecosistema emprendedor [9].

La matriz DOFA ampliada se diferencia de la DOFA clásica en la profundidad del análisis realizado y en el aporte de estrategias para el contexto estudiado. Mientras que la matriz DOFA se enfoca en identificar y describir factores internos (debilidades y fortalezas) y externos (oportunidades y amenazas), la DOFA

ampliada combina los factores internos y externos para definir estrategias cruzadas, más específicas y detalladas [9]. Ejemplos de las combinaciones entre factores son:

- *Fortalezas - Oportunidades (FO)*: Estrategias para aprovechar las fortalezas y oportunidades.
- *Fortalezas - Amenazas (FA)*: Estrategias para usar las fortalezas y minimizar las amenazas.
- *Debilidades - Oportunidades (DO)*: Estrategias para superar las debilidades aprovechando las oportunidades.
- *Debilidades - Amenazas (DA)*: Estrategias defensivas para minimizar las debilidades y evitar las amenazas.

4. MATRÍZ DOFA DEL ECOSISTEMA EMPRENDEDOR VENEZOLANO

Utilizando la información recopilada del análisis de las variables ecosistémicas, se identificaron las principales debilidades, oportunidades, fortalezas y amenazas del ecosistema emprendedor venezolano, respecto a la innovación y la tecnología. La Tabla II resume las ideas principales de estos elementos en una matriz DOFA.

Tabla II: Elementos Clave de la Matriz DOFA del Ecosistema Emprendedor en Venezuela

Factores Internos	
Debilidades	Fortalezas
D1. Emprendimientos de subsistencia o por necesidad.	F1. Formación técnica o universitaria en el 33% de quienes emprenden.
D2. Débil interrelación de cooperación intersectorial.	F2. Plataforma de Ciencia y Tecnología para la Innovación (PCTI).
D3. Deserción o abandono educativo.	F3. Percepción positiva entre emprendedores que propicia la creación de redes.
D4. Demanda inestable y bajo poder adquisitivo.	
D5. Bajo nivel de madurez tecnológica.	
D6. Ausencia de marco regulatorio claro para el emprendimiento de base tecnológica y científica.	
Factores Externos	
Oportunidades	Amenazas
O1. Iniciativas locales exitosas que promueven la innovación y el emprendimiento.	A1. Ambiente de inestabilidad y de desconfianza.
O2. Interés en la transformación digital y en la necesidad de actualización tecnológica.	A2. Fuga de talentos por migración.
O3. Aborde del emprendimiento desde los primeros niveles educativos.	A3. Política de importaciones abiertas con cero o muy bajo arancel.
O4. Alternativas nacionales e internacionales de financiamiento al emprendimiento.	A4. Factores internos que desestiman la demanda.
O5. Espacios de encuentro entre sectores.	A5. Posibilidad de que se establezcan nuevas medidas sancionatorias al país.
	A6. Falta o insuficientes opciones de financiamiento o fuentes de crédito.

Estos elementos clave observados en la revisión bibliográfica, agrupados en la matriz DOFA en factores internos y factores externos, se describen con mayor detalle a continuación:

4.1 Debilidades

- D1. 96% de los emprendimientos que surgen entran en la

categoría de emprendimientos de subsistencia o por necesidad, con poca o nula innovación y bajo nivel de tecnología [1].

- D2. PCTI con una débil interrelación de cooperación intersectorial público, privado y académico, una de las razones es la falta de alineación Empresa – Academia [10].
- D3. Deserción o abandono educativo, así como un alto nivel de inasistencia, brecha que se amplía en los siguientes niveles educativos [11].
- D4. Condiciones de la demanda inestables y bajo poder adquisitivo [12].
- D5. Bajo nivel de madurez tecnológica empresarial [13].
- D6. La ausencia de un marco regulatorio claro dirigido al emprendimiento tecnológico [14].

4.2 Oportunidades

- O1. Existen iniciativas locales exitosas que buscan promover la innovación y el emprendimiento, lo que puede impulsar el desarrollo del ecosistema [15].
- O2. El interés en la transformación digital, así como la conciencia en la necesidad de la actualización tecnológica, que abre la puerta al desarrollo de algunas innovaciones tecnológicas [13].
- O3. Posibilidad de abordar el emprendimiento desde los primeros niveles de educación [10].
- O4. Existen alternativas para el financiamiento y asesoría de organismos internacionales para fortalecer el ecosistema emprendedor [1].
- O5. Necesidad de espacios de encuentro entre los sectores los cuales sirven para propiciar la cooperación intersectorial [15].

4.3 Fortalezas

- F1. 33% de quienes emprenden actualmente tienen formación técnica o universitaria [1].
- F2. Existe una Plataforma de Ciencia y Tecnología para la Innovación (PCTI) que, aunque relativamente reciente, ya dispone de 24.000 investigadores y cuenta con capacidades de investigación académica con potencial para la producción de tecnología [10].
- F3. Percepción positiva del emprendedor y la emprendedora venezolana que facilita la construcción de redes de apoyo, cooperación y consumo [1].

4.4 Amenazas

- A1. El ambiente o la percepción de inestabilidad y desconfianza en las instituciones, influyen en el deterioro de las condiciones económicas del país y compromete el éxito de las iniciativas de cooperación y financiamiento [1].
- A2. La migración acelerada ha provocado la fuga de talentos, especialmente de profesionales en niveles técnico y universitario [11][16].
- A3. La política de importaciones “abiertas”, con cero (0) o bajos aranceles perjudica grandemente la producción nacional y la hace poco competitiva [16].
- A4. La emocionalidad postelectoral, la sensación de

pérdida de libertades, la autocensura, el bajo nivel adquisitivo de parte de la población, entre otros factores, desestimulan la demanda [17].

- A5. Posibilidad de que se establezcan nuevas medidas sancionatorias al país o directamente sobre alguna de sus autoridades de gobierno [18].
- A6. Falta de financiamiento y modalidades de crédito privado [1].

5. MATRÍZ DOFA AMPLIADA DEL ECOSISTEMA EMPRENDEDOR VENEZOLANO

Luego de identificar los factores internos y externos más relevantes del ecosistema emprendedor y su impacto positivo o negativo en el emprendimiento, el siguiente proceso en esta investigación ha sido analizar las relaciones o sinergias entre estos factores, con el objetivo de alinear estrategias y aprovechar las oportunidades que ofrece este ecosistema en Venezuela.

Esto permitió elaborar estrategias dinamizadoras, reactivas, adaptativas y de supervivencia que integran un mapa estratégico que evidencia la situación actual y las áreas de abordaje clave para potenciar el desarrollo de emprendimientos de base científica y tecnológica en el país.

5.1 Estrategias Dinamizadoras (Fortalezas-Oportunidades)

- *Potenciar el Resultado de las Iniciativas Locales Exitosas (combina: F2 - O1)*

El programa *EmprendeTech* [15] realizado por el Centro de Innovación y Emprendimiento del IESA bajo el auspicio de BID Lab ofrece la oportunidad de formación a jóvenes emprendedores de base tecnológica e impacto social. Esta puede mencionarse como una iniciativa exitosa porque en su corto tiempo de desarrollo exhibe excelentes resultados.

Replicar esta iniciativa en otros centros de educación superior e invitar a científicos(as) e investigadores(as) a participar activamente, con sus trabajos de investigación puede, además de ampliar la cobertura de estas actividades, impulsar resultados que conducen a incrementar la producción de tecnología.

Bajo esta estrategia dinamizadora sería interesante impulsar la reforma de la Ley Orgánica de Ciencia, Tecnología e Innovación (LOCTI), en su artículo 41, con el fin de introducir la posibilidad de canalizar parte de los aportes hacia el financiamiento de emprendimientos tecnológicos surgidos a partir de las actividades de investigación, desarrollo e innovación (I+D+i); los cuales son promovidos por la misma ley. Se requiere la participación de todos, por un lado el Estado y por el otro el Sector Privado como grupo de interés, posible comprador, licenciatario o consumidor de tecnología, apoyado además en su necesidad consciente de actualización tecnológica.

- *Crear una Plataforma Colaborativa para la Innovación con las Agrupaciones Industriales (F2 - O2)*

Desde la cooperación y la urgencia de soluciones comunes a problemas básicos como pueden ser la necesidad de nuevos materiales, mejores procesos o incluso la actualización tecnológica, pueden agruparse organizaciones para impulsar estrategias de innovación cómo por ejemplo la innovación

abierta que se puede implementar de la mano de las Universidades que son, en Venezuela, los mayores productores de investigación.

5.2 Estrategias Reactivas (Fortalezas - Amenazas)

- *Introducir Innovaciones Tecnológicas en los Emprendimientos de Subsistencia (combina: F1 - A2, A6)*

Fomentar la creación de soluciones innovadoras con uso tecnológico para los emprendimientos de subsistencia o necesidad. Por ejemplo, creando servicios de apoyo o de gestión de clientes. Para esto se requiere de incentivos directos que sirvan para desarrollar estas soluciones y estructurarlas dentro de un marco de transformación digital que además funcione como palanca de crecimiento de estos emprendimientos.

Esta iniciativa busca reducir el impacto negativo de las amenazas presentes con el apoyo de las fortalezas presentes.

5.3 Estrategias Adaptativas (Debilidades - Oportunidades)

- *Crear Espacios de Alineación Empresa - Academia (D2, D4, D5 - O2, O5)*

Son necesarios espacios donde empresarios puedan manifestar a los investigadores, las diferentes problemáticas que existen en el sector. En este mismo espacio, los investigadores podrían exhibir las posibles soluciones que han diseñado o están diseñando para el sector. Esto, además de facilitar la alineación, también fortalecerá el espacio de oportunidades al robustecer la PCTI y podría mejorar el nivel de madurez tecnológico empresarial.

Como parte de estos espacios de conversación, es necesario se definan estructuras de apoyo que faciliten la vinculación, por ejemplo: con mecanismos de Protección Intelectual, con la estrategia legal apropiada, entre otras áreas de apoyo. Aquí es posible que las incubadoras o aceleradoras de negocio, como parte del ecosistema, puedan estar interesadas en participar.

5.4 Estrategias de Supervivencia (Debilidad - Amenazas)

- *Crear Vías Posibles para el Financiamiento y el Levantamiento de Capital (D1 - A1, A6)*

Implementar espacios de vinculación entre inversores, interesados y emprendedores tecnológicos, con el objetivo de: mejorar las habilidades de levantamiento de capital, acercar los emprendedores a la inversión y crear oportunidades de financiamiento.

La naturaleza reactiva y de supervivencia de esta estrategia la posiciona en un lugar destacado dentro de la jerarquía estratégica, indicando su importancia relativa frente a otras iniciativas.

6. RESULTADOS

El ecosistema emprendedor de base científica y tecnológica en Venezuela es un sistema adaptativo complejo, por lo que ésta es un área de investigación que continúa en desarrollo. Sin embargo, el análisis de las condiciones del ecosistema emprendedor del país realizadas en el marco de esta investigación permite presentar tres resultados clave:

1. Análisis del ecosistema emprendedor tecnológico venezolano:

Se realizó un análisis estratégico del ecosistema emprendedor

tecnológico en Venezuela considerando factores internos y externos. Se utilizó el modelo de visión sistémica del ecosistema emprendedor desarrollado por Kantis y Federico (2020), el cual contiene tres ejes (Capital humano, Espacio de oportunidades, e Interacciones y redes) y diez dimensiones para evaluar la situación de los emprendimientos tecnológicos.

2. Matriz DOFA ampliada:

Se aplicó una matriz DOFA ampliada para identificar puntos de intervención y mejora para el ecosistema emprendedor. La matriz DOFA ampliada diseñada en la investigación combina el análisis de factores internos (debilidades y fortalezas) y externos (oportunidades y amenazas) para definir estrategias cruzadas más detalladas que las aportadas en una matriz DOFA clásica.

3. Estrategias para impulsar el ecosistema emprendedor tecnológico:

A partir del análisis DOFA ampliado, se formuló un conjunto de estrategias para impulsar el desarrollo de emprendimientos de base científica y tecnológica (EBCT) en Venezuela. Las estrategias propuestas buscan conformar un marco de referencia para fomentar el EBCT a fin de impulsar la creación de empresas de alto impacto y acelerar el surgimiento de "unicornios" o *startups* que logren alcanzar valoraciones superiores a mil millones de dólares.

7. CONCLUSIONES Y RECOMENDACIONES

El ecosistema emprendedor tecnológico venezolano posee tanto fortalezas como debilidades significativas que influyen directamente en su capacidad para aprovechar oportunidades y superar los desafíos que presenta el entorno. Es alentador observar la percepción positiva del emprendedor venezolano y la existencia de una Plataforma de Ciencia y Tecnología para la Innovación (PCTI). Sin embargo, la alta proporción de emprendimientos de subsistencia con poca innovación y tecnología, así como la débil interrelación de cooperación intersectorial dentro de la PCTI, representan desafíos importantes.

En el contexto del país es importante priorizar las acciones de supervivencia. Urge encontrar soluciones para apalancar financieramente a los emprendimientos tecnológicos, considerando entre otras cosas, el costo y el tiempo dentro de cualquier proyecto de desarrollo tecnológico.

Destacan sin embargo, fortalezas y oportunidades que pueden usarse como puntos de apoyo para impulsar cambios en el ecosistema. Es el caso de la cantidad de investigadores registrados en la PCTI, el interés empresarial en la transformación digital y la existencia de iniciativas locales exitosas. Sobre estas se puede pivotar para el diseño de muchas otras estrategias ofensivas dirigidas directamente al crecimiento del ecosistema.

Es necesario complementar las políticas públicas, dirigidas a la creación de medios de vida y por tanto al emprendimiento de subsistencia o necesidad, con acciones que estructuralmente impulsen el desarrollo país. Esta necesidad de complementariedad podría cubrirse con el diseño e implementación de políticas públicas específicamente dirigidas al desarrollo del emprendimiento científico tecnológico.

Los espacios de encuentro, comunicación y acuerdo trisectorial

son fundamentales para la implementación de las acciones estratégicas, aún más si consideramos que la mayoría de ellas representan acciones sencillas cuya implementación depende, casi totalmente, de la voluntad de cooperar en pro de objetivos comunes.

La implementación de estrategias que fomenten la colaboración entre el sector académico y el empresarial, así como la creación de plataformas de financiamiento, son fundamentales para el éxito de los emprendimientos de base tecnológica en el país. La reforma de la LOCTI y la promoción de espacios de alineación entre empresas y academia son pasos necesarios para fortalecer la innovación y la competitividad. Al unir esfuerzos entre el Estado, el sector privado y las instituciones educativas, se puede construir un entorno propicio para el crecimiento sostenible de los emprendimientos, garantizando así un futuro más próspero y tecnológicamente avanzado para la sociedad venezolana.

Como recomendación, consideramos que resultaría particularmente útil para fomentar el desarrollo del emprendimiento científico tecnológico en el país, el inventariar quiénes son las actoras y los actores del ecosistema que están interesados en la creación de tecnología y en trabajar estableciendo mecanismos de interconexión. Del mismo modo es importante conocer las fortalezas y potencialidades de su oferta de investigación, servicios y productos, ya que ayudaría a focalizar la inversión y a maximizar el resultado de los proyectos de apoyo al emprendimiento y de los espacios de vinculación. Finalmente valdría la pena explorar alternativas innovadoras de capitalización y financiamiento dentro del ecosistema.

REFERENCIAS

- [1] A. Freitez, P. Monteferrante, A. Rodríguez, D. Marotta, E. Ojeda, G. García, M. Ponce, y L. Lauriño, *Informe GEM 2022-2023. Emprendimiento en Venezuela: Desafíos, Transformaciones y Oportunidades*, 2023, <https://elucabista.com/wp-content/uploads/2024/04/Infome-GEM-Vzla-2024-Final.pdf>
- [2] H. Kantis and J. Federico, *A Dynamic Model of Entrepreneurial Ecosystems Evolution*, *J. Evol. Stud. Bus.*, vol. 5, no. 1, pp. 182-220, January 2020, <https://doi.org/10.1344/jesb2020.1.j072>
- [3] H. Kantis y P. Angelelli, *Emprendimientos de Base Científico-Tecnológica en América Latina. Importancia, Desafíos y Recomendaciones para el Futuro*, Banco Interamericano de Desarrollo (BID), February 2020, https://publications.iadb.org/publications/spanish/document/Emprendimientos_de_base_cientifico-tecnologica_en_America_Latina_Importancia_desafios_y_recomendaciones_para_el_futuro.pdf
- [4] M. Morris, X. Neumeier, Y. Jang, and D. Kuratko, *Distinguishing Types of Entrepreneurial Ventures: An Identity-Based Perspective*, *J. Small Bus. Manage.*, vol. 56, no. 3, pp. 453-474, 2018. DOI: 10.1111/jsbm.12272
- [5] G. Hidalgo, M. Kamiya, and M. Reyes, *Nro 16. Emprendimientos Dinámicos en América Latina. Avances en Prácticas y Políticas*, Serie Políticas Públicas y Transformación Productiva, no. 16, Corporación Andina de Fomento (CAF), Caracas, Venezuela 2014, <https://scioteca.caf.com/handle/123456789/371>
- [6] J. Federico, S. Ibarra, y H. Kantis, *La Visión Sistémica y el Ecosistema Emprendedor: Herramientas para la Enseñanza Universitaria*, Prodem, España, Septiembre 2019, <https://prodem.ungs.edu.ar/la-vision-sistemica-del-emprendimiento-y-el-ecosistema-emprendedor>
- [7] D. Iseberg, *Applying the Ecosystem Metaphor to Entrepreneurship: Uses and Abuses*, *The Antitrust Bulletin*, vol. 61, no. 4, pp. 564-573, December 2016, DOI: 10.1177/0003603X16676162
- [8] H. Kantis, J. Federico, and S. Ibarra, *Condiciones Sistémicas para el Emprendimiento Dinámico. Las Brechas Abiertas de América Latina*:

- ¿Convergencia o Divergencia?*, Informe Prodem, 1ra. ed., Asociación Civil Red Pymes Mercosur, Octubre 2018, <https://prodem.ungs.edu.ar/wp-content/uploads/2018/10/Informe-PRODEM-2018.pdf>
- [9] H. Ponce, *La Matriz FODA: Una Alternativa para Realizar Diagnósticos y Determinar Estrategias de Intervención en las Organizaciones Productivas y Sociales*, Revista Contribuciones a la Economía, ISSN 16968360, Septiembre 2006, <https://eco.mdp.edu.ar/cendocu/repositorio/00290.pdf>
- [10] Ministerio de Ciencia y Tecnología, *Plan Nacional de Ciencia, Tecnología e Innovación. Construyendo un Futuro Sustentable: Venezuela 2005-2030*, Ministerio de Ciencia y Tecnología, Caraca, Venezuela, Octubre 2005, <https://www.cenditel.gob.ve/portal/wp-content/uploads/biblioteca/2005/pncti/pncti.pdf>
- [11] UCAB, *Encuesta Nacional de Condiciones de Vida (ENCOVI) 2023. Radiografía de la Vulnerabilidad Social en Venezuela y Propuestas de Políticas Públicas*, Universidad Católica Andrés Bello, Caracas, Venezuela, Marzo 2023, <https://dialnet.unirioja.es/servlet/articulo?codigo=9596498>
- [12] CONINDUSTRIA, *Encuesta Coyuntura Industrial I Trimestre 2024*, Confederación Venezolana de Industriales, Mayo 2024, <https://conindustria.org/encuestas-de-coyuntura-industrial>
- [13] C. Alvarez y L. Díaz, *Chequeo Digital en Venezuela: Resultados y Reflexiones sobre la Madurez Digital de las Empresas en un Entorno Económico Desafiante*, Banco Interamericano de Desarrollo (BID), Marzo 2023, <http://dx.doi.org/10.18235/0004827>
- [14] Gaceta Oficial, *Ley para el Fomento y Desarrollo de Nuevos Emprendimientos*, G.O. No. 6.656 Extraordinaria, Octubre 2021, <https://sapi.gob.ve/wp-content/uploads/2024/03/ley-para-el-fomento-y-desarrollo-de-nuevos-emprendimientos-20211116202828.pdf>
- [15] IESA, *17 Emprendimientos Miden el Éxito de sus Propuestas de Valor Ante Jurados de EmprendeTECH*, Instituto de Estudios Superiores de Administración (IESA), Febrero 2024, <https://www.iesa.edu.ve/iesa-al-dia/17-emprendimientos-miden-el-exito-de-sus-propuestas-de-valor-ante-jurados-de-emprendetech>
- [16] Fedecámaras, *Camino al Futuro Venezuela 2025. Bases para la Discusión y Construcción Colectiva de Un Nuevo Modelo de Desarrollo*, 2022, <https://www.fedecamaras.org.ve/wp-content/uploads/2022/07/Camino-al-futuro.pdf>
- [17] M. Díaz, *Tensión Poselectoral Genera Incertidumbre en los Venezolanos*, Diario La Verdad, Agosto 23, 2024, <https://laverdad.com/tension-poselectoral-genera-incertidumbre-en-los-venezolanos>
- [18] Bloomberg News, *EE.UU. Prepara Nuevas Sanciones Contra Funcionarios Venezolanos Tras Elecciones*, Bloomberg, Septiembre 03, 2024, <https://www.bloomberg.com/news/articles/2024-09-03/ee-uu-prepara-sanciones-contra-venezolanos-tras-dudosa-eleccion>

Importancia de la Catalogación Digital en la Conservación del Patrimonio Natural: Ficoflora Venezuela, Un Caso de Estudio

Yusneyi Carballo-Barrera¹, Santiago Gómez-Acevedo²
yusneyicarballo@gmail.com, chachacho@gmail.com

¹ Escuela de Computación, Universidad Central de Venezuela, Caracas, Venezuela

² Instituto de Biología Experimental (IBE), Universidad Central de Venezuela, Caracas, Venezuela

Resumen: La catalogación digital es crucial en la preservación del patrimonio, ya que, al aprovechar las tecnologías digitales es posible garantizar la longevidad y la accesibilidad a materiales patrimoniales valiosos, contribuyendo a su salvaguarda y posibilitando su apreciación como un legado. Centrándonos en el marco de esta investigación, la catalogación digital es un componente esencial para la preservación y difusión de información sobre biodiversidad, contribuyendo al monitoreo efectivo, la conservación y la gestión sostenible de los ecosistemas. En el contexto de la ficoflora de Venezuela, este artículo de revisión documental examina la relevancia de las bases de datos y de los metadatos en la conservación de la diversidad biológica, facilitando la gestión y el acceso a información sobre especies de macroalgas bentónicas marinas, así como los trabajos de investigación y localidades donde se reportan. Primero, se destaca la importancia de los marcos de trabajo y de los estándares de catalogación para asegurar integridad, accesibilidad e interoperabilidad de metadatos y datos, promoviendo la colaboración, la visibilidad de la información y el uso sostenible de la biodiversidad a partir del conocimiento y preservación de sus datos. Luego, se analizan las metodologías empleadas para la sistematización de colecciones biológicas, la digitalización de registros georreferenciados y la creación de catálogos accesibles en línea, describiendo como caso de estudio el Catálogo Digital Ficoflora Venezuela, cuyo objetivo es documentar la diversidad de la ficoflora marina, a la vez que proporciona datos actualizados y con calidad científica para la investigación, la conservación y la educación ambiental. Este trabajo contribuye a la discusión sobre la necesidad de integrar enfoques innovadores en la gestión del patrimonio natural, considerando a la catalogación digital como una herramienta para documentar y proteger recursos tangibles e intangibles, tanto culturales como naturales.

Palabras clave: Ficoflora Venezuela; Catalogación digital; Preservación patrimonio natural; Metadatos de biodiversidad; *Marine algae biodiversity data*; *Digital preservation of biodiversity information*; *Open access databases for biodiversity*.

1. INTRODUCCIÓN

La preservación del patrimonio natural es una tarea crucial en el contexto actual de cambio ambiental y pérdida de biodiversidad. Este artículo de revisión documental se estructura en varias secciones que abordan la importancia de la catalogación digital para la conservación de datos asociados a la ficoflora mundial, y particularmente, la presente en las costas de Venezuela.

La estructura del artículo incluye una sección que detalla los métodos y materiales utilizados para crear catálogos digitales de biodiversidad en general, para luego avanzar a la descripción del caso de estudio: el Catálogo Taxonómico Digital de la Ficoflora de Macroalgas Bentónicas Marinas de Venezuela.

A lo largo del artículo se abordan conceptos, procesos y enfoques presentes en la catalogación digital, como la sistematización de colecciones biológicas, la gestión de datos digitales, la digitalización de registros georreferenciados y la creación de bases de datos accesibles en línea. También se especifican cuáles son los principales marcos de trabajo y estándares reseñados en la bibliografía para la definición de metadatos, la creación de bases de datos de biodiversidad y de catálogos de colecciones naturales.

Para el catálogo digital Ficoflora Venezuela la metodología empleada se basa en la recopilación y análisis de datos sobre la

diversidad de macroalgas bentónicas marinas en las costas de Venezuela, utilizando herramientas digitales para la catalogación y el almacenamiento de información. Se describen cuáles son las fases del proceso de digitalización de la información y la catalogación de registros requeridas para crear un repositorio accesible y útil, tanto para investigadores como para el público en general, a fin de facilitar la toma de decisiones informadas para la gestión y conservación de la ficoflora marina.

El desarrollo de este artículo tiene como objetivo contribuir al entendimiento de la catalogación digital como un recurso esencial para la conservación del patrimonio natural, promoviendo la generación de datos de acceso abierto, la colaboración entre instituciones y el uso sostenible de la biodiversidad marina.

2. LA CATALOGACIÓN DIGITAL COMO HERRAMIENTA PARA LA PRESERVACIÓN DEL PATRIMONIO

En el contexto de la preservación del patrimonio natural y de la biodiversidad marina, la catalogación digital se refiere al proceso de identificación, recopilación y almacenamiento digital de datos, imágenes y registros sobre inventarios de especies, ecosistemas y recursos naturales con el objetivo de asegurar la accesibilidad a esta información y la conservación a largo plazo, a partir de datos basados en conocimiento científico que son objetivos, fiables y comparables [1][2].

Este proceso facilita la toma de decisiones informada, el monitoreo de cambios y la implementación de estrategias efectivas para preservar la biodiversidad. Se apoya en bases de datos y aplicaciones digitales que son diseñadas para contener información detallada sobre especies, hábitats, reportes, ocurrencias, usos, amenazas y medidas de conservación, por lo que es esencial para la actual gestión sostenible de los ecosistemas y sus recursos [3].

En un sentido más amplio, la preservación digital permite que el patrimonio tangible e intangible se representen en formatos que garantizan su longevidad y accesibilidad en el presente y para las generaciones futuras. Lo anterior implica el establecimiento de archivos digitales o de modelos digitales (en 2D y 3D) con los cuales se registra y se conserva información útil para documentar, proteger y dar a conocer especies, artefactos o sitios valiosos [4][5].

El desarrollo de un proyecto de catalogación digital contempla varios conceptos, procesos y enfoques cuyos elementos característicos se indican en la Figura 1.

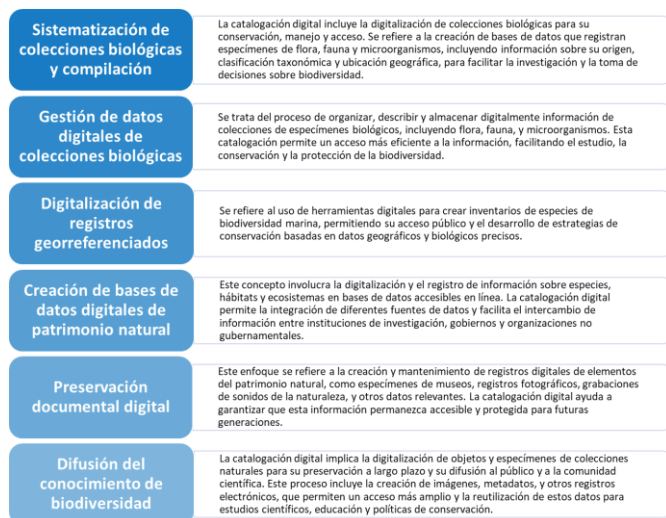


Figura 1: Conceptos, Procesos y Enfoques Presentes en la Catalogación Digital

La “Carta de la UNESCO sobre la Preservación del Patrimonio Digital” destaca la importancia de los proyectos de digitalización para salvaguardar el patrimonio [6]. Las reconstrucciones digitales, las representaciones virtuales, los gemelos digitales y los sistemas informativos integrados introducen metodologías innovadoras para estudiar, preservar y difundir, ofreciendo perspectivas en donde se integran enfoques tradicionales y nuevos, además de herramientas para la conservación y la valorización del patrimonio cultural y natural [7][8][9].

El desarrollo de catálogos digitales es esencial para la conservación de la biodiversidad marina debido a que aporta información de utilidad para:

- Diseñar líneas base, líneas de tiempo o estudios comparativos que permiten analizar la resiliencia de los ecosistemas biológicos frente a perturbaciones naturales o inducidas por acciones humanas [10].

- Conocer la diversidad de los ecosistemas marinos, estimar su valor y definir políticas de protección de áreas marinas protegidas las cuales son cruciales para preservar la biodiversidad y respaldar los servicios ecosistémicos [11].
- Comprender la distribución y abundancia de especies maradoras o monitoras, por ejemplo, la distribución y proliferación de algas marinas como marcadores de impactos naturales o antropogénicos en los entornos marinos [12].
- Descubrir nuevos compuestos bioactivos con propiedades farmacológicas, médicas (con acción anticoagulante, antioxidante o antiproliferativa), alimenticias o industriales, los cuales demuestran la infinita ventana de oportunidades actuales y futuras de la biodiversidad marina, incluidas las macroalgas [13].
- Diseñar proyectos de investigación en las áreas de genética, ecología, etnobotánica, manejo turístico, acuicultura o ecoproducción que requieren información ambiental y de la biodiversidad marina, incluida la ficoflora [14].

La preservación digital de la información asociada con la diversidad biológica o biodiversidad es crucial para monitorear y conservar componentes vitales de los ecosistemas marinos. Hernández [15, p. 17] resalta que la biodiversidad “*constituye per se patrimonio natural vital para la supervivencia de la especie humana. La diversidad de especies, los genes que ellas contienen y los ecosistemas donde viven, son parte importante del patrimonio natural*” ya que de esta diversidad provienen poblaciones de plantas, animales, microorganismos, sustratos y ecosistemas fundamentales para un desarrollo sostenible y el mantenimiento de la vida sobre el planeta.

3. BASES DE DATOS, METADATOS Y ESTÁNDARES ABIERTOS PARA EL REGISTRO DE LA FICOFLORA

El avance de la informática aplicada al registro y la preservación de la biodiversidad, facilitado por las tecnologías de Internet, permite la gestión de grandes cantidades de datos sobre ficoflora o macroalgas marinas (*marine algae biodiversity data*).

Varias son las plataformas utilizadas para este fin. Algunas permiten visualizar y difundir información sobre la diversidad de grupos específicos, lo que contribuye al estudio y los esfuerzos de conservación por ejemplo de las macroalgas bentónicas marinas o de los invertebrados marinos, por dar dos ejemplos concretos [16][17][18]. Otras plataformas buscan recopilar grandes conjuntos de datos (*datasets*) que se registran en bases de datos de biodiversidad de acceso abierto (*open access databases for biodiversity*), con mapas digitales o atlas en línea de regiones específicas, proporcionando herramientas valiosas para comprender y conservar la biodiversidad marina única de áreas como el Triángulo de Coral en el Indo-Pacífico o las costas del mar Caribe [19][20].

3.1. Metadatos de Registros de Biodiversidad

La creación de bases de datos de patrimonio natural involucra la digitalización y el registro de información sobre especies, hábitats y ecosistemas en bases de datos accesibles en línea. Son bases de datos que integran diferentes fuentes de datos y facilitan el intercambio de información entre instituciones de

investigación, gobiernos y organizaciones, además de recopilar información de utilidad para profesionales en distintas áreas, docentes, estudiantes y el público en general [16].

Para organizar, catalogar y registrar la biodiversidad marina, y en particular de las macroalgas bentónicas, punto focal de nuestros proyectos de investigación, se utilizan bases de datos y también documentos con fuentes bibliográficas, fotografías, mapas, e incluso representaciones digitales de muestras físicas que están en museos naturales o en herbarios. Esta diversidad de recursos requiere de la definición de un conjunto de datos descriptivos de mayor nivel de abstracción denominados metadatos.

Los metadatos, en el contexto de preservación de la biodiversidad, son información estructurada que describe, explica y contextualiza los datos sobre biodiversidad, facilitando su gestión, uso y conservación a largo plazo. Abarcan información sobre descripción de las metodologías usadas para coleccionar los datos, el rango de los datos, fechas, alcance geográfico y taxonómico [21]. Como un ejemplo puntual, los metadatos del *dataset* “*Benthic Macroalgae From the Coasts of Venezuela 1822-2021*” incorporados en el *Caribbean OBIS Node (Ocean Biodiversity Information System)* [20][22][23] se especifican indicando (Figura 2):

- **Title:** Título para el recurso.
- **Type:** Tipo de recurso, contemplando cuatro opciones:
 - Metadatos del recurso (*Resource metadata*).
 - Listas taxonómicas (*Checklist data*).
 - Registros biológicos (Occurrence data).
 - Eventos de muestreo (*Sampling event data*).
- **Publication date:** Fecha de publicación.
- **Metadata last modified:** Fecha de la última actualización de la metadata y de los datos.
- **Project ID:** Código de identificación del proyecto.
- **Hosted by:** Organización publicadora, por ejemplo, *Caribbean OBIS Node*.
- **Network:** Red, por ejemplo OBIS, la cual trabaja en colaboración con GBIF (*Global Biodiversity Information Facility*) y la IOC (*Intergovernmental Oceanographic Commission*).
- **DOI o Digital Object Identifier:** Identificador de Objeto Digital. El DOI es el identificador de preferencia, aunque se pueden indicar otros alternativos.
- **Description:** Descripción del conjunto de datos.
- **Purpose:** Propósito de la investigación o proyecto de recopilación y registro de datos.
- **Temporal scope:** Alcance, ámbito o rango temporal.
- **Geographic scope:** Alcance geográfico y georreferenciación.
- **Taxonomic scope:** Alcance taxonómico.
- **Methodology:** Metodología de muestreo, de recopilación, de procesamiento, de transformación, de estandarización y/o de mapeo de los datos.

- **Additional info:** Información adicional.
- **Bibliography:** Bibliografía utilizada en la especificación de algunos campos del recurso, en particular en la descripción o metodología.
- **Contacts:** Contactos de las investigadoras o los investigadores responsables del proyecto de colección de los datos.
- **Administrative, Technical and Installation contacts:** contacto de la persona(s) encargada de la gestión del proyecto o nodo en el cual se incluye el conjunto de datos.
- **Endpoints:** Puntos de enlace a archivos de especificación de metadatos y de datos en formatos estandarizados e interoperables [24].
- **Endorsed by:** Entidad avaladora, por ejemplo, *Ocean Biodiversity Information System*.
- **Data description language:** Lenguaje en el cual se expresan los metadatos y la data.
- **Citation:** Cita sugerida para el recurso.



Figura 2: Vista Parcial de la Metadata del Conjunto de Datos “*Benthic Macroalgae From the Coasts of Venezuela 1822-2021*”

Los metadatos de preservación son fundamentales para garantizar que los valiosos datos sobre biodiversidad permanezcan accesibles, comprensibles y utilizables, contribuyendo así a los esfuerzos globales de conservación y gestión sostenible de la diversidad biológica. En la Tabla I se precisan doce (12) aspectos que reflejan la relevancia de los metadatos [21][25][26]:

Tabla I: Aspectos que Definen la Relevancia de los Metadatos para la Preservación e Intercambio de Información de Biodiversidad

Aspecto	Relevancia de los Metadatos
1. Interoperabilidad	Contribuyen a la interoperabilidad entre diferentes sistemas y archivos digitales, facilitando el intercambio y la integración de datos de biodiversidad, lo cual es esencial para compartir información entre los elementos involucrados en la conservación (investigación, gestión de áreas protegidas y toma de decisiones políticas).
2. Contextualización	Proporcionan el contexto necesario para interpretar correctamente los datos sobre especies, ecosistemas y recursos genéticos, incluyendo el análisis temporal y espacial.
3. Calidad y Confiabilidad	Los metadatos aumentan la confiabilidad y la calidad de la información al documentar el origen, los métodos de recolección y el procesamiento de los datos de biodiversidad.

Aspecto	Relevancia de los Metadatos
4. Valor científico	Aumentan el valor científico de los datos al permitir su verificación, replicación y reutilización en nuevos estudios e investigaciones sobre biodiversidad.
5. Uso eficaz de la información	Los metadatos permiten una gestión más eficiente de los flujos de datos sobre biodiversidad, lo que lleva a una toma de decisiones más efectiva para definir y apoyar políticas de conservación.
6. Colaboración	Los metadatos, así como las bases de datos, que se estructuran siguiendo estándares mejoran la interoperabilidad y el intercambio de información de utilidad para: <ul style="list-style-type: none"> • Coordinar esfuerzos de conservación entre múltiples organizaciones. • Tomar decisiones informadas y con una visión más completa del estado de la biodiversidad. • Priorizar iniciativas de conservación de manera colaborativa y consensuada.
7. Evaluación de Impacto Ambiental (EIA)	En el contexto de la EIA, los metadatos ayudan a comprender, interpretar y utilizar eficazmente la información sobre la presencia de especies en el espacio y el tiempo, lo cual es necesario para respaldar esfuerzos de mitigación, restauración o compensación de impactos sobre la biodiversidad.
8. Cobertura y Representatividad	Los metadatos asociados a las áreas protegidas y las bases de datos de biodiversidad permiten: <ul style="list-style-type: none"> • Analizar la cobertura de los sistemas de áreas protegidas. • Identificar vacíos en la protección de especies o ecosistemas clave. • Priorizar la creación o expansión de áreas protegidas en zonas subrepresentadas.
9. Identificación de amenazas	Los metadatos pueden incluir información sobre factores de amenaza para la biodiversidad la cual se utiliza para: <ul style="list-style-type: none"> • Identificar las principales presiones sobre especies y ecosistemas. • Priorizar acciones de conservación dirigidas a mitigar amenazas específicas. • Desarrollar estrategias de conservación más efectivas y focalizadas.
10. Priorización y Conservación	Al incluir información sobre fechas, ubicaciones geográficas y contexto los metadatos permiten realizar análisis para: <ul style="list-style-type: none"> • Identificar tendencias temporales o la distribución espacial de la biodiversidad. • Detectar cambios en los ecosistemas a lo largo del tiempo. • Priorizar áreas que requieren intervención inmediata o con alto valor de conservación.
11. Toma de decisiones	Apoyan la toma de decisiones informada en la conservación de la biodiversidad al proporcionar información detallada sobre la procedencia y calidad de los datos.
12. Gestión a largo plazo	Permiten la administración efectiva de colecciones digitales de datos sobre biodiversidad, asegurando su accesibilidad y usabilidad en el futuro.

3.2. Bases de Datos y Acceso Abierto a Información de Diversidad Biológica

Las bases de datos de biodiversidad son herramientas y, a la vez, recursos digitales que contribuyen significativamente a la

investigación y la conservación de la mismas, aportando datos primarios para muchas regiones y grupos taxonómicos [27]. Las bases de datos desempeñan un papel vital en la identificación de taxones, georreferenciación de la ubicación de especies, especificación de fuentes de trabajos de investigación, participantes y de proyectos.

Las bases de datos son una herramienta de prioridad para la preservación, registrando y documentando la biodiversidad de manera eficaz [28], suministran información de utilidad tanto para el registro de las especies, como para el descubrimiento o identificación de especies nuevas, especies raras o de identidad taxonómica incierta [29].

Para diseñar bases de datos de acceso abierto para la preservación de la biodiversidad de la ficoflora es fundamental considerar marcos de trabajo (*frameworks*) de catalogación y estándares que aseguren integridad, cobertura, interoperabilidad y reducción de sesgos en los esfuerzos taxonómicos. Estos marcos y estándares son valiosos porque permiten:

- Establecer un marco abierto para el diseño de bases de datos de biodiversidad, facilitando también su vinculación.
- Compartir y reutilizar metadatos, datos, documentos y recursos digitales.
- Desarrollar ciencia ciudadana que conduzca a un monitoreo más amplio, al apoyo de estudios macroecológicos y a los esfuerzos de conservación.
- Proporcionar información accesible para la conservación y la gestión de la biodiversidad.

En la Tabla II se identifican doce (12) marcos de trabajo (MT) y estándares de interoperabilidad (E) que son utilizados para la creación de bases de datos de uso general y, especialmente, para crear bases de datos específicas interoperables y de acceso abierto para registros de biodiversidad marina [30-41].

Tabla II: Marcos de Trabajo (MT) y Estándares (E) de Interoperabilidad para la Definición de Metadatos y la Creación de Bases de Datos de Biodiversidad

Nombre	Tecnología	Descripción	Ventajas y Limitaciones
OData (<i>Open Data Protocol</i>) (E)	REST ¹ JSON Estándares ISO/IEC: 20802:1, 20801:2, y 20922	Protocolo abierto para consultar y actualizar datos. Creado por Microsoft en 2007.	V: Facilita la creación y consumo de APIs con arquitectura REST. Compatible con múltiples plataformas. L: Puede ser complejo para implementaciones simples.
Dublin Core (E)	XML RDF	Conjunto de elementos de metadatos para describir recursos	V: Ampliamente adoptado. Flexible y extensible. L: Puede ser

¹ Significado de las siglas: REST (*Representational State Transfer*), JSON (*JavaScript Object Notation*), JSON-LD (*JSON for Linking Data*), ISO/IEC (*International Organization for Standardization / International Electrotechnical Commission*), XML (*Extensible Markup Language*), RDF

(*Resource Description Framework*), RDFa (*Resource Description Framework in Attributes*), NetCDF (*Network Common Data Form*), LSID (*Life Science Identifier*), API (*Application Programming Interface*), OGC (*Open Geospatial Consortium*), EML (*Ecological Metadata Language*).

Nombre	Tecnología	Descripción	Ventajas y Limitaciones
		digitales. Desarrollado por DCMI (<i>Dublin Core Metadata Initiative</i>) en 1995.	demasiado general para descripciones detalladas.
Schema.org (E)	JSON-LD RDFa	Vocabulario estructurado para marcar contenido web. Fundado por Google, Microsoft, Yahoo y Yandex en 2011.	V: Mejora la visibilidad en motores de búsqueda. Amplia cobertura de dominios. L: Puede ser complejo para implementar correctamente.
Darwin Core (E)	XML RDF	Estándar para compartir datos de biodiversidad. Desarrollado como Estándar de Información de Biodiversidad (<i>Biodiversity Information Standards</i>) por TDWG (<i>Taxonomic Data Working Group</i>) en 1998.	V: Ampliamente utilizado en biodiversidad. Facilita la integración de datos. L: Puede ser limitado para datos muy especializados.
OBIS-ENV-DATA (<i>OBIS with Environmental Data</i>) (MT)	Darwin Core NetCDF	Extensión de Darwin Core para datos ambientales marinos. Desarrollado por OBIS en 2016.	V: Permite vincular datos biológicos y ambientales. L: Requiere conocimiento de Darwin Core.
EML (<i>Ecological Metadata Language</i>) (E)	XML	Estándar de metadatos para datos ecológicos. Desarrollado por NCEAS (<i>National Center for Ecological Analysis and Synthesis</i>) en 1997.	V: Detallado y específico para ecología. L: Puede ser complejo para usuarios no especializados.
WoRMS (<i>World Register of Marine Species</i>) (MT)	LSID REST API	Registro taxonómico mundial de especies marinas. Lanzado en 2008 a partir del Registro Europeo de Especies Marinas.	V: Proporciona identificadores únicos para especies. L: Enfocado principalmente en taxonomía.
GBIF IPT (<i>Integrated Publishing Toolkit</i>) (MT)	Darwin Core EML	Herramienta para publicar datos de biodiversidad. Desarrollado por GBIF en 2010.	V: Facilita la publicación y compartición de datos. L: Requiere cierta experiencia técnica para su configuración.
<i>Ocean Data Standards</i> (E)	NetCDF OGC	Estándares para intercambio de datos oceanográficos. Desarrollado por el IODE (<i>International</i>	V: Ampliamente adoptado en oceanografía. L: Puede ser complejo para implementaciones pequeñas.

Nombre	Tecnología	Descripción	Ventajas y Limitaciones
		<i>Oceanographic Data and Information Exchange</i> de la UNESCO.	
OAI-PMH (<i>Open Archives Initiative - Protocol for Metadata Harvesting</i>) (MT)	Dublin Core XML	Protocolo para recolección y transmisión de metadatos sobre cualquier soporte electrónico y codificados en Dublin Core. Desarrollado por la OAI (<i>Open Archives Initiative</i>) en 2001.	V: Ampliamente utilizado en repositorios. L: Limitado a metadatos.
ABCD (<i>Access to Biological Collection Data</i>) (E)	XML	Acceso a datos de colecciones biológicas. Desarrollado por TDWG en 2005.	V: Detallado para colecciones biológicas. L: Puede ser complejo de implementar.
MiXS (<i>Minimum Information about Any (x) Sequence</i>) (E)	XML	Información mínima sobre cualquier secuencia genética. Desarrollado por GSC (<i>Genomic Standards Consortium</i>) en 2011.	V: Estandariza metadatos genómicos. L: Requiere conocimientos genómicos.

3.3. Estándares para la Catalogación de Colecciones Biológicas

La representación de los datos de biodiversidad a través del *Access to Biodiversity Collections Data* (ABCD) o del Darwin Core, entre otros estándares, es esencial para un eficaz intercambio e integración de datos asociados la observación de organismos, su taxonomía, localización, contexto ecológico o geológico. Los estándares cuentan al menos con un conjunto de elementos básicos o características que son aplicables en muchos dominios biológicos (por ejemplo la botánica), pero también pueden integrar extensiones hacia subdisciplinas específicas (por ejemplo, listas de verificación taxonómica, herbarios, recursos genéticos), además de establecer colaboraciones con otras organizaciones para adaptar o ampliar un estándar hacia nuevas disciplinas, como la genómica, la metagenómica o las secuencias de marcadores genéticos [42].

Iniciativas como *Integrated Publishing Toolkit* (IPT) o el estándar de datos de descriptores de colecciones *TDWG Latimer Core* abordan los desafíos en el intercambio de datos de biodiversidad a la vez que promueven la ciencia abierta y la movilización de datos digitales [43].

La catalogación de colecciones naturales representa un desafío dada la diversidad de partes interesadas y sus enfoques, incluyendo a investigadoras(es), proyectos, instituciones y patrocinantes, además de usuarios con distintos niveles o necesidades. Sin embargo, se reconoce que la catalogación digital de colecciones biológicas es fundamental para [42]:

- Crear mapas de recursos con muestras que respaldan la investigación taxonómica.
- Localizar y contactar a personas o a instituciones que poseen colecciones y especímenes.
- Destacar la importancia de los museos y de los herbarios.
- Determinar vacíos o ausencia de datos sobre los especímenes en las colecciones.
- Construir un gráfico de conocimiento sobre la biodiversidad en una localidad, área o región.
- Incentivar la definición, financiamiento y desarrollo de proyectos de digitalización y preservación de colecciones.

Hobern *et al.* [44][45] indica que a los catálogos de colecciones de biodiversidad se les puede dar varios usos:

1. Directorio con siglas institucionales y de colecciones estandarizadas.
2. Localización de especímenes y de datos genéticos.
3. Insumo para la creación de bases de datos de colecciones, incluyendo también conjuntos de datos, imágenes, publicaciones, información de expediciones, notas de campo, entre otros recursos de información.
4. Estimación del tamaño, cobertura, brechas y valor de las colecciones.
5. Vincular los datos de los especímenes en las colecciones con publicaciones, proyectos o autoridades externas.
6. Especificar, validar, mejorar y estandarizar métodos de conservación de las colecciones.
7. Reducir la duplicación de esfuerzos de recopilación y curación de datos creando colecciones reutilizables.
8. Hacer más accesible la información de las colecciones, llegando con nuevos enfoques y servicios a un público más amplio, a nuevos usuarios, a agregadores de datos, comunidades de investigadores(as), entes financiadores, agencias gubernamentales o a responsables de definir políticas públicas.
9. Incrementar la citación, atribución y visibilidad de las colecciones, además del vínculo entre sus datos.
10. Comprender cómo la singularidad de su información otorga a las colecciones de biodiversidad un estatus de recurso irremplazable, por lo que la colaboración nacional, continental o global es una vía para integrar esfuerzos y recursos que permitan su incremento, mantenimiento, digitalización, divulgación, entre otras actividades.
11. Generar métricas y obtener indicadores (crecimiento de la colección, calidad de sus datos, utilidad de las especies identificadas, entre otros) que puedan presentarse a partes interesadas en apoyar o financiar el mantenimiento de los catálogos.

Como ya se indicó, la creación de catálogos de colecciones es un desafío porque se necesita lograr una unificación de enfoques, alcanzar una visión y unos objetivos compartidos, definir metadatos y datos interoperables, entre otros requisitos.

Aunque no es un trabajo fácil, en la Tabla III se identifican, a partir de la traducción y adaptación [44], varios estándares de datos de relevancia que ayudan a definir catálogos de

colecciones con datos biológicos, paleontológicos, mineralógicos, geológicos, multimedia, además de otros tipos de elementos físicos o digitales.

Tabla III: Estándares de Datos para Crear Catálogos de Colecciones

Estándar	Descripción
1. Darwin Core (DwC)	Es el estándar más utilizado para compartir datos sobre especímenes de historia natural y observaciones de biodiversidad. Se basa en estándares de metadatos existentes, entre otros el Dublin Core y cuenta con el respaldo repositorios de datos a nivel de especímenes, herramientas o plataformas que colaboran con el <i>Taxonomic Data Working Group (TDWG)</i> .
2. ABCD	Es un estándar alternativo para datos de especímenes que suministra un esquema de catalogación de datos de colecciones biológicas. Se le considera un estándar integral, completo y estructurado para datos de biodiversidad.
3. ABCDEF G	El <i>Access to Biological Collection Databases Extended for Geosciences</i> es una extensión de ABCD desarrollada para catalogar y preservar datos paleontológicos, mineralógicos y geológicos en colecciones digitalizadas.
4. Audubon Core (AC)	Es un conjunto de vocabularios diseñados para representar metadatos para recursos multimedia de biodiversidad (imágenes, audios, videos), cuadernos de campo o las colecciones de dichos recursos.
5. TDWG Collection Descriptions (Latimer Core)	El estándar de datos de descripciones de colecciones de TDWG Latimer Core define un conjunto de clases y propiedades útiles para representar grupos de objetos de colección y su información asociada. Estos incorporan características utilizadas para: <ul style="list-style-type: none"> • Describir, agrupar y desglosar colecciones. • Definir métricas para cuantificar las colecciones. • Crear identificadores persistentes para rastrear colecciones y sus contrapartes digitales. El modelo <i>Collection Descriptions</i> de TDWG es el sucesor natural o la continuación del estándar <i>Natural Collections Descriptions (NCD)</i> .

En este punto, es importante destacar el trabajo del *TDWG Attribution Project*, el cual más que un estándar, es una colaboración entre TDWG y *Research Data Alliance* para mejorar los estándares existentes o crear nuevos, otorgando atribución por el mantenimiento, la conservación y la digitalización de objetos físicos y digitales, con un énfasis especial en las colecciones de biodiversidad. Dos resultados a destacar de esta colaboración son [44][46]:

1. Incorporación de actualizaciones en el estándar Darwin Core y sus plantillas.
2. La actualización del Estándar para la Descripción de Colecciones o *TDWG Collection Descriptions Data Standard* el cual es una herramienta en evolución que busca proporcionar un marco coherente para la descripción de colecciones naturales, mejorando la interoperabilidad entre diferentes recursos de datos, facilitando la agregación y el intercambio de información sobre colecciones naturales y catálogos.

Con el objetivo de facilitar la búsqueda, la identificación, la recuperación y la reutilización de metadatos y datos en colecciones, catálogos, conjuntos de datos o bases de datos, se han establecido los principios FAIR (*Findable, Accessible, Interoperable and Reusable*). Los principios rectores FAIR

(FAIR Guiding Principles) identifican cuatro atributos que hacen posible el descubrimiento, la recuperación y el intercambio de conocimiento por parte de humanos y de máquinas a través de algoritmos y aplicaciones [47][48]:

a) *Findable* (Encontrables)

Para que los datos sean encontrables, deben tener: a) Identificadores únicos, globales y persistentes para todos los recursos de información, sean estos metadatos, listas taxonómicas, registros biológicos o eventos de muestreo. Se recomienda asignar un Identificador de Objeto Digital o DOI a los metadatos y los conjuntos de datos; b) Descriptores a través de metadatos enriquecidos; c) Registro e indexación en recursos.

b) *Accessible* (Accesibles)

Para que los datos sean accesibles: a) Se deben poder recuperar mediante su identificador usando un protocolo estandarizado; b) El protocolo debe ser abierto, gratuito y universalmente implementable; c) Se puede requerir autenticación y autorización cuando sea necesario; d) Los metadatos deben ser accesibles incluso cuando los datos ya no estén disponibles.

c) *Interoperable* (Interoperables)

Para que los datos sean interoperables: a) Deben usar un lenguaje formal, accesible, compartido y ampliamente aplicable; b) Utilizar vocabularios que sigan los principios FAIR; c) Incluir referencias calificadas a otros metadatos.

d) *Reusable* (Reutilizables)

Para que los datos sean reutilizables: a) Deben estar bien descritos con múltiples atributos precisos y relevantes; b) Publicarse con una licencia de uso de datos clara y accesible; c) Asociarse con un fuente de origen o información de proveniencia detallada; d) Cumplir con los estándares relevantes de la comunidad.

Los principios FAIR son parte integral del trabajo en ciencia abierta, describiendo pautas para una adecuada gestión de data, con acceso abierto a metadatos y datos de investigación.

En el contexto de la preservación de la biodiversidad de las macroalgas, el uso de estándares, de metadatos y de bases de datos tienen varias aplicaciones, más aún cuando se utilizan para generar publicaciones divulgativas o se incorporan en aplicaciones informáticas que hagan posible el acceso público a la información de colecciones biológicas mediante Internet:

- Listas de verificación (*checklists*) de macroalgas marinas que pueden mejorar la sistemática y la integración de estudios sobre su biodiversidad y claves taxonómicas (*taxonomic keys, identification keys*) para facilitar el proceso de identificación en cualquier nivel de la jerarquía taxonómica [49][50][51, p. 18-27][52].
- Publicaciones de libros, catálogos de especies o guías ilustradas que contribuyen al conocimiento de la biodiversidad en general, o son un aporte a la actualización del conocimiento de un grupo particular, por ejemplo, la ficoflora de macroalgas bénticas en una región [14][29].
- Bases de datos o conjuntos de datos (*datasets*) de con

información sobre ficoflora y con acceso en línea que desempeñan un papel importante en la organización y el intercambio de datos de biodiversidad con la comunidad científica y el público en general [16][17][22][53].

- Bancos de datos genéticos que se utilizan para la gestión, la conservación y la trazabilidad de biodiversidad marina [54] o también, para corroborar las identificaciones morfológicas de especies [55][56].
- Iniciativas como *Ocean Biodiversity Informatics* (OBI) que progresivamente migró hacia el *Ocean Biodiversity Information System* (OBIS) o el *Integrated Publishing Toolkit* (IPT) que proporcionan herramientas dinámicas e integrales para estudiar la biodiversidad marina o su distribución geográfica [34][43][37].
- Bases de datos de biodiversidad integrales, de alcance geográfico amplio y de acceso abierto que ayudan a visibilizar data de registros de especies y también a identificar las lagunas en los registros, facilitando la preservación de la biodiversidad como es el *Caribbean OBIS Node* o el *Global Registry of Scientific Collections* (GRSciColl) [20][57].

Al adherirse o cumplir con los estándares de catalogación establecidos, aprovechar las bases de datos de biodiversidad de acceso abierto, promover la estandarización, la digitalización de datos y de colecciones, el intercambio de datos y las estrategias de catalogación colaborativa, se incrementa la eficiencia y la eficacia de las iniciativas de preservación de la biodiversidad de las algas marinas, respaldando así los esfuerzos de investigación, conservación y gestión.

4. FICOFLORA VENEZUELA: CATÁLOGO DIGITAL DE ACCESO ABIERTO PARA PRESERVACIÓN DE PATRIMONIO NATURAL

El proyecto Catálogo Taxonómico Digital de la Ficoflora de Macroalgas Bentónicas Marinas de Venezuela, en adelante Ficoflora Venezuela, tiene como objetivo principal preservar y divulgar, en formatos de libre acceso, la información de reportes y registros de macroalgas bentónicas publicada en la bibliografía venezolana e internacional [16].

La costa venezolana presenta una importante diversidad de macroalgas bentónicas marinas, la cual ha sido estudiada y documentada mediante diversos proyectos, sin embargo, se observan situaciones que han dificultado el acceso a estos datos:

- Especies poco documentadas, con identidad taxonómica incierta, raras o distribución geográfica restringida.
- Numerosos trabajos de investigación y referencias bibliográficas que contienen reportes de registros de la ficoflora venezolana no tienen la visibilidad adecuada o exigen una consulta presencial.
- Se carece de datos geográficos de las poblaciones naturales o hay referencias a las localidades de los reportes, pero no se especifican las coordenadas geográficas de su ubicación.
- Los datos están desactualizados o dispersos.
- Se dispone de listas ficoflorísticas, sin descripciones, claves, ilustraciones, ni mapas, haciendo de estas listas herramientas poco prácticas y con bajo atractivo didáctico.

En consecuencia, resulta necesario el levantamiento y sistematización de la información existente en una base de datos estandarizada, eficiente y disponible para consulta web o móvil.

4.1. Importancia de los Catálogos Digitales de Ficoflora

Venezuela posee una importante información florística y taxonómica, pero son datos dispersos que se encuentran en bibliotecas, revistas, libros, informes, trabajos especiales de grado y tesis de postgrado, proyectos, herbarios, entre otras fuentes. En consecuencia, resulta necesario mantener un trabajo constante de revisión de fuentes de información, recopilación de datos, verificación, actualización de cambios en las denominaciones taxonómicas válidas, corrección en algunos casos de impresiones o errores, además de la estandarización de los datos, su registro en bases de datos y la definición de metadatos que los describan [29].

El desarrollo y mantenimiento de catálogos digitales de biodiversidad, como es el caso de Ficoflora Venezuela, es importante porque son herramientas fundamentales para:

a) *Posibilitar el acceso a investigaciones sin barreras geográficas o de horario:* Los catálogos digitales son fuente de información de calidad, verificada y constantemente actualizada, con datos útiles para el desarrollo de la docencia y de proyectos de investigación, conservación o productivos.

b) *Divulgar e incrementar la visibilidad:* Los catálogos digitales en línea dan visibilidad a los proyectos que se desarrollan y ayudan a posicionar en el ecosistema mundial de investigación a quienes los realizan, sus identificaciones, reportes y descubrimiento de nuevas especies.

c) *Preservar en formatos digitales el conocimiento:* La UNESCO ha destacado que los soportes informáticos que permiten el almacenamiento y el respaldo del patrimonio digital, representan “recursos de valor perdurable dignos de ser conservados para las generaciones futuras” [6].

Ficoflora Venezuela constituye así un respaldo mediante tecnologías digitales de trabajos de investigación, catálogos, reportes de campo, notas de investigación o colecciones de exiccatas, incrementando el patrimonio digital del país y ayudando a la preservación de información de biodiversidad que se encuentra, en muchos casos, en formatos analógicos.

d) *Generar datos abiertos y apoyar las metas de desarrollo sostenible:* A partir de proyectos como el Catálogo Taxonómico Digital Ficoflora Venezuela se pueden derivar conjuntos de datos de acceso abierto (*open datasets*) [22]. También suministran información que puede ayudar con el alcance de los Objetivos de Desarrollo Sostenible (ODS) de la Agenda 2030 [58] mediante la creación de recursos de información, materiales educativos y servicios de consulta digital.

e) *Suministrar datos de calidad para alimentar tecnologías habilitadoras del acceso al conocimiento:* Los datos almacenados en los catálogos taxonómicos digitales permiten generar información codificada para etiquetas de identificación, taxonomía y descripción de especies (en formato de código de barras o de QR, por ejemplo), suministran datos de georreferenciación de reportes que pueden ser consultados en mapas de acceso abierto, además de permitir la consulta de

especies, investigaciones o localidades en aplicaciones que pueden ser utilizadas desde múltiples dispositivos, en cualquier momento y desde cualquier lugar.

4.2. Proceso de Digitalización en Ficoflora Venezuela

El proyecto Catálogo Digital Ficoflora Venezuela ha permitido registrar reportes de colecciones de algas marinas presentados en trabajos académicos y científicos de distinta naturaleza, publicados desde 1822 hasta el presente. Esta iniciativa se ha desarrollado en distintas fases, iniciando hace más de 35 años con la compilación de reportes de ficoflora en registros almacenados, inicialmente, en hojas de cálculo y utilizando posteriormente gestores de bases de datos relacionales para crear repositorios con entidades y campos normalizados. La creación del Catálogo Taxonómico Nacional de las Macroalgas Bentónicas Marinas de Venezuela, con su base de datos y aplicación web, fue un proyecto desarrollado entre 2012 y 2016 [29]. Desde el 2016 hasta la fecha, se ha desarrollado un esfuerzo de incremento de los registros de reportes, actualización del catálogo e inclusión de herramientas en la aplicación web que permite el acceso a información ficoflorística. El ciclo de vida del proceso de digitalización del proyecto Ficoflora Venezuela tiene siete fases, brevemente descritas en la Figura 3.

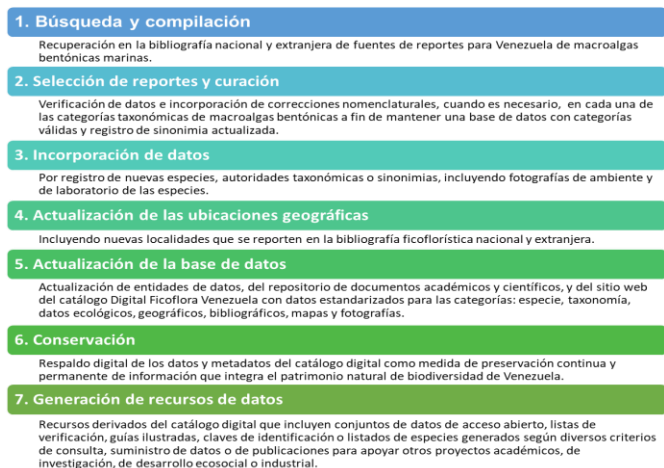


Figura 3: Fases del Ciclo de Vida del Proceso de Digitalización en el Proyecto Ficoflora Venezuela

4.3. Estructura de Metadatos, Base de Datos y Arquitectura de la Aplicación Web del Catálogo Ficoflora Venezuela

El Catálogo Digital Ficoflora Venezuela tiene como núcleo una base de datos de registros de reportes de especies de macroalgas marinas distribuidas en toda la costa continental y de las Dependencias Federales del país. Estos datos se almacenan en cuatro grandes categorías de entidades: 1. Taxonomía, 2. Referencias bibliográficas, 3. Registro de reportes, 4. Ubicación geográfica. En el diagrama de la Figura 4 se observa cómo se agrupan las entidades tablas en cada categoría.

Las entidades de datos preservan la información necesaria para responder a las consultas realizadas en la aplicación Ficoflora Venezuela, según cuatro criterios principales de búsqueda: por especie, por clasificación taxonómica, por ubicación o por referencia bibliográfica. Las consultas generan distintos listados

que pueden paginarse e imprimirse, además de contar con filtros para hacer búsquedas más específicas. La aplicación utiliza MySQL como sistema de gestión de bases de datos [59].

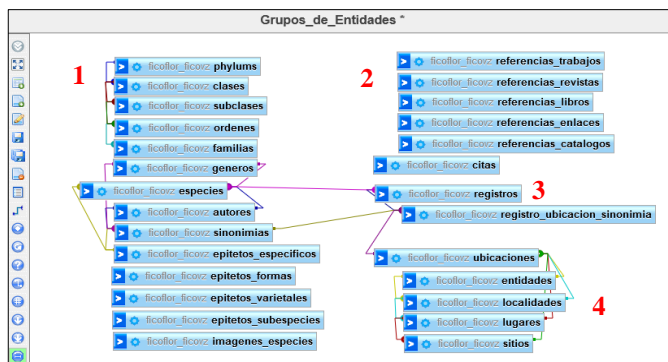


Figura 4: Categorías de Entidades de Datos en el Catálogo Digital Ficoflora Venezuela

De la base de datos se extraen campos que se combinan en una ficha descriptiva para cada especie, en donde se incluye el nombre válido, clasificación taxonómica, sinonimia, fotografías (hábito, detalles morfológicos, detalles anatómicos), reportes, cita a las autoras o los autores de los reportes, la referencia bibliográfica en detalle de las publicaciones, además de mapas con puntos de ubicación georreferenciados (ver Figura 5).



Figura 5: Vista Parcial de la Ficha de Datos de una Especie

En la ficha hay controles que permiten ajustar la cantidad de registros a mostrar en cada página del listado de resultados e incluye filtros que permiten buscar por cualquier texto o elemento de información contenido en la sección de reportes o en la sección de referencias bibliográficas, por lo que es muy útil para ajustar las búsquedas para un autor(a), por año, ubicación, nombre o tipo de publicación, entre otros criterios.

Los mapas generados en el Catálogo Digital Ficoflora Venezuela muestran el punto de ubicación de cada reporte identificando el nombre de la localidad y sus coordenadas (Figura 6), mapas que se puede imprimir, con escalas que se puede aumentar o reducir. Estos mapas son creados utilizando librerías de código abierto de JavaScript, específicamente con librerías y APIs de Leaflet y OpenStreetMap [60][61].



Figura 6: Mapa y Puntos Georreferenciados de Reportes de Especies

4.4. Tecnologías, Módulos y Arquitectura de Software de la Aplicación Catálogo Digital Ficoflora Venezuela

La aplicación web Ficoflora Venezuela es una interfaz que da acceso público a la base de datos y al catálogo digital de macroalgas bénticas de Venezuela y está dirigida a cualquier persona interesada en consultar su inventario, con reportes georreferenciados y actualizados, beneficiando a ficólogos, estudiantes, docentes, investigadores, alguicultores y comunidades interesadas en el desarrollo de modelos ecológicos de producción de algas marinas.

a) Metodología y Tecnologías

La aplicación Ficoflora Venezuela [59][60] se desarrolló siguiendo las fases y lineamientos de la metodología de Desarrollo Adaptable de Software, tiene un diseño adaptativo (*responsive*) que permite su visualización adecuada en móviles y otras pantallas, con interfaces creadas utilizando librerías de código abierto de Bootstrap con códigos basados en HTML, CSS y JavaScript. Para el desarrollo de la aplicación se utilizó el *framework* Laravel, con un patrón de arquitectura de software MVC (Modelo-Vista-Controlador), un motor de gestión de plantillas Blade que permite crear vistas reutilizables, Eloquent para el mapeo objeto-relacional (ORM, *Object-Relational Mapping*) y la gestión de la interacción con la base de datos MySQL. Para la creación de gráficos estadísticos de los reportes incluidos en el catálogo digital se utiliza C3.js. La exportación de listados de datos, de fichas de especies y de mapas en formato de documento portable se realiza mediante la librería mPDF, mientras que para la importación y exportación de plantillas de datos en formato de hoja de cálculo se utiliza la API Laravel-Excel. La aplicación se ejecuta bajo una configuración de Servidor Virtual Privado (VPS, *Virtual Private Services*) con sistema operativo CentOS 8 Stream 8.

b) Módulos y Arquitectura

El Catálogo Digital Ficoflora Venezuela está compuesto por dos módulos (ver Figura 7):

1. Módulo Gestión de Datos y Generación de Estadísticas: Este módulo contiene dos submódulos. El submódulo Gestor de Datos tiene como objetivo principal facilitar la gestión de los datos de los reportes de especies de macroalgas que alimentan e incrementan la base de datos del catálogo digital, permitiendo

que personas que no pertenecen al área de la computación puedan contribuir de manera sencilla y segura a incorporar nuevos datos. El submódulo Generador de Estadísticas es el otro componente, permitiendo la consulta de información cuantitativa consolidada: básicamente cantidades de registros de especies, taxonomías, bibliografías o ubicaciones geográficas que se han incorporado al catálogo digital.

2. *Módulo de Consulta*: Tiene como objetivo servir de interfaz a la información recopilada en la base de datos de reportes de especies, además de suministrar información general sobre el proyecto de investigación Ficoflora Venezuela, el catálogo digital, las opciones de consulta, contactos y créditos.

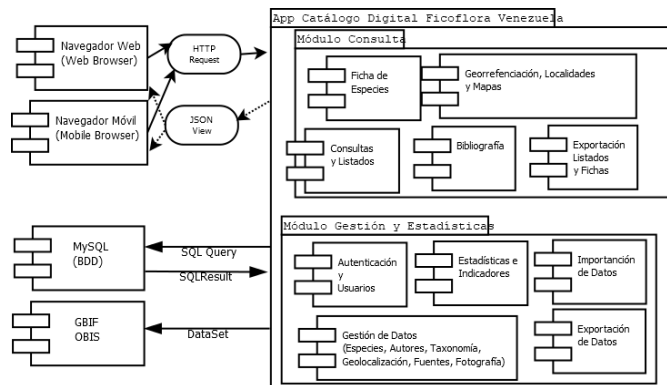


Figura 7: Componentes e Interacciones de la Aplicación Ficoflora Venezuela

4.5. Mapeo de Metadatos del Catálogo Ficoflora Venezuela con Darwin Core y OBIS-GBIF

En marco del Proyecto “*Rescuing the Knowledge Base of Venezuela’s Marine Biodiversity*” (BID-CA2020-025-NAC, 2021-2022) [61] la data del catálogo digital Ficoflora Venezuela se integró al Nodo Caribe en la plataforma OBIS-GBIF mediante el conjunto de datos titulado “*Benthic Macroalgae From the Coasts of Venezuela 1822-2021*”, aportando el mayor número de datos de ocurrencia, con 10.438 registros.

La taxonomía de las especies reportadas en la base de datos Ficoflora Venezuela fue previamente validada utilizando el *World Register of Marine Species* y fue estandarizada bajo Darwin Core, con un formato de datos DwC & GBIF API, en codificación UTF-8 *character encoding*. La correspondencia entre las entidades y campos de la base de datos Ficoflora Venezuela, con los metadatos y datos considerados por parte del GBIF, se realizó respetando el vocabulario de términos definido por Darwin Core a través del mapeo de campos utilizando los recursos del IPT de GBIF: *Metadata, Event, Occurrence, Citations, Multimedia, Verbatim, y Rights*.

5. RESULTADOS

Las algas marinas han constituido un importante recurso biológico para muchos países a lo largo de muchas décadas, lo que ha generado la necesidad de enriquecer el conocimiento taxonómico de este grupo vegetal para así contribuir a su aprovechamiento y conservación.

El Catálogo Digital Ficoflora Venezuela ha documentado la presencia de 705 especies (incluidas categorías infraespecíficas), 257 referencias bibliográficas, 6.237 registros, distribuidos en 11.895 ubicaciones y 574 localidades que abarcan toda la costa del territorio nacional.

Los conjuntos de datos generados, a partir de la base de datos Ficoflora Venezuela y siguiendo estándares de catalogación internacionales como el Darwin Core, alimentan sistemas de información que amplían la visibilidad y el acceso de la comunidad científica global a información de biodiversidad, cumpliendo así con los principios de posibilidad de consulta y recuperación, accesibilidad, interoperabilidad y reusabilidad (FAIR). Iniciativas de este tipo representan un avance significativo para el conocimiento y la investigación de la ficoflora, con importantes implicaciones para la conservación de su biodiversidad y desarrollo sostenible.

6. CONCLUSIONES

La catalogación de registros de reportes de macroalgas marinas, su divulgación y su visibilización aporta información de utilidad para la evaluación de las cuencas hidrográficas, hidrogeológicas y de las zonas marino-costeras de Venezuela. Inventariando estas algas como un recurso hidrobiológico se genera información de valor para su aprovechamiento sustentable.

El Catálogo Taxonómico Digital de la Ficoflora de Macroalgas Bentónicas Marinas de Venezuela también aporta datos con los cuales se pueden identificar algunas especies que producen sustancias tóxicas, ocasionan graves daños directos o indirectos a la salud, son marcadores de alteraciones ambientales o se constituyen, por su proliferación, en elementos dañinos para el ecosistema y para los organismos que lo compartimos. Este catálogo contribuye al desarrollo del conocimiento científico y tecnológico al permitir recopilar en una base de datos de acceso abierto una gran variedad de datos útiles para la conservación y preservación de las macroalgas bentónicas marinas como parte del patrimonio nacional y mundial de biodiversidad.

Mediante la catalogación digital se preserva en el ciberespacio un conocimiento científico útil para reconocer la distribución de nuestra riqueza natural, aumentar la sensibilización ambiental y la protección de la biodiversidad marina como un valor inestimable de la nación.

AGRADECIMIENTOS

El desarrollo de los proyectos enmarcados en esta investigación ha sido posible gracias al patrocinio o al financiamiento de:

- El Consejo de Desarrollo Científico y Humanístico de la Universidad Central de Venezuela (CDCH-UCV), a través del Proyecto de Grupo PG 03-8643-2013.
- El Fondo Nacional de Ciencia, Tecnología e Innovación (FONACIT) y el Ministerio de Ciencia y Tecnología, a través de los proyectos PEI- FONACIT-2011001216 y FONACIT 2023PGP247.
- CaribeSur, Proyecto BID-CA2020-025-NAC.
- Apoyo institucional y logístico de la Coordinación de Investigación y de la Coordinación Administrativa de la

Facultad de Ciencias; del Instituto de Biología Experimental, del Centro de Enseñanza Asistida por Computador y del Centro de Computación, Facultad de Ciencias, Universidad Central de Venezuela; del Instituto Experimental Jardín Botánico Dr. Tobías Lasser; del Instituto Pedagógico de Miranda José Manuel Siso Martínez; y del Centro de Investigaciones Ecológicas Guayacán, Núcleo Sucre de la Universidad de Oriente.

REFERENCIAS

- [1] Gobierno de España. *Estrategia de Biodiversidad y Ciencia (2023-2027)*, Ministerio de Ciencia e Innovación, 2023, <https://fundacion-biodiversidad.es/wp-content/uploads/2022/12/ESTRATEGIA-DE-BIODIVERSIDAD-Y-CIENCIA-2023-2027.pdf>.
- [2] Gobierno de España. *Inventario Español del Patrimonio Natural y la Biodiversidad*, Ministerio para la Transición Ecológica y el Reto Demográfico, s.f., <https://www.miteco.gob.es/es/biodiversidad/temas/inventarios-nacionales/inventario-espanol-patrimonio-natural-biodiv.html>
- [3] S. Choy, N. Crofts, R. Fisher, N. Lek Choh, S. Nickel, C. Oury, and K. Šlaska. *Directrices UNESCO/PERSIST sobre Selección del Patrimonio Digital para su Conservación a Largo Plazo*, International Journal of Engineering and Advanced Technology, Marzo 2016, https://repository.ifa.org/bitstream/123456789/1220/1/persist-content-guidelines_es.pdf
- [4] A. Ahmad and S. Sharma, S. *Sustainable Digital Preservation and Access of Heritage Knowledge in India*. Desidoc Journal of Library & Information Technology, vol. 40, no. 5, pp. 321-325, Noviembre 2020, <https://doi.org/10.14429/djlit.40.05.15822>
- [5] M. Casas-Pérez, F. Castillo-Rivera, G. Moreno-Chavez, and D. Sarocchi, *Fringe Projection Method's Evolving Role in Cultural Asset Cataloging and Conservation: Potentiality in Testing Conservative Restoration Techniques*, Journal of Anthropological and Archaeological Sciences, vol. 9, no. 3, Abril 2024, <https://doi.org/10.32474/jaas.2024.09.000313>
- [6] UNESCO.org, *Noción de Patrimonio Digital*, UNESCO Collection, 2021, <https://web.archive.org/web/20230612012520/https://es.unesco.org/themes/information-preservation/digital-heritage/concept-digital-heritage>
- [7] A. Gil, *Digital Reconstructions - A Methodology for the Study, Preservation and Dissemination of Architectural Heritage*, Proceedings of the 8th International Congress on Archaeology, Computer Graphics, Cultural Heritage and Innovation ARQUEOLÓGICA 2.0, Valencia, Spain, September 5-7, 2016.
- [8] C. Zou, S-Y. Rhee, L. He, D. Chen, and X. Yang, *Sounds of History: A Digital Twin Approach to Musical Heritage Preservation in Virtual Museums*, Electronics, April 2024, vol. 13, no. 12, 2388, <https://doi.org/10.3390/electronics13122388>
- [9] A. Marra, S. Gerbino, A. Greco, and G. Fabbrocino, *Combining Integrated Informative System and Historical Digital Twin for Maintenance and Preservation of Artistic Assets*, Sensors, vol. 21, no. 17, 5956, June 2021, <https://doi.org/10.3390/s21175956>
- [10] D. Fautin, P. Dalton, L. Incze, J. Leong, C. Pautzke, A. Rosenberg, P. Sandifer, G. Sedberry, J. Tunnell Jr, I. Abbott, R. Brainard, M. Brodeur, L. Eldredge, M. Feldman, F. Moretzsohn, P. Vroom, M. Wainstein, and N. Wolff, *An Overview of Marine Biodiversity in United States Waters*, Plos One, vol. 5, no. 8, e11914, August 2010, <https://doi.org/10.1371/journal.pone.0011914>
- [11] M. Rodríguez-Ruiz, M. Andreu-Cazenave, C. Ruz, C. Ruano-Chamorro, F. Ramírez, C. González, S. Carrasco, A. Pérez-Matus, and M. Fernández, *Initial Assessment of Coastal Benthic Communities in the Marine Parks at Robinson Crusoe Island*, Latin American Journal of Aquatic Research, vol. 42, no. 4, pp. 918-936, August 2014, <https://doi.org/10.3856/vol42-issue4-fulltext-16>
- [12] S. Rouse, J. Loxton, M. Jones, and J. Porter, *A Checklist of Marine Bryozoan Taxa in Scottish Sea Regions*. Zookeys, 787, pp. 135-149, October 2018, <https://doi.org/10.3897/zookeys.787.24647>
- [13] L. Araújo, J. Silva, B. Aguiar, J. Bezerra, A. Rodrigues, and F. Paulini, *Antiproliferative Activity of Marine Brown Algae-Derived Compounds: A Review*, World Journal of Advanced Research and Reviews, vol. 11, no. 1, pp. 060-072, July 2011, <https://doi.org/10.30574/wjarr.2021.11.1.0306>
- [14] S. Gómez-Acevedo, M. García-Ortiz, Y. Carballo-Barrera, y N. Gil-Luna, *Macroalgas Bénticas del Parque Nacional Archipiélago Los Roques Venezuela. Guía Ilustrada*, EdiCiencias UCV, Noviembre 2017, https://www.researchgate.net/publication/377014313_Macroalgas_Bénticas_del_Parque_Nacional_Archipielago_Los_Roques_Venezuela_Guia_Ilustrada
- [15] E. Hernández, *Programa de Desarrollo de Capacidades para el Caribe para el Patrimonio Mundial. Módulo 6 Gestión del Patrimonio Natural*, Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO), Mayo 2011.
- [16] Web Ficoflora Venezuela, *Catálogo Digital de la Ficoflora de Venezuela*, publicación electrónica, Universidad Central de Venezuela, Caracas, Editores: Santiago Gómez, Yusneyi Carballo Barrera, Mayra García y Nelson Gil Luna, Agosto 2024, <https://www.ficofloravenezuela.info.ve/public/index>
- [17] Y. Carballo-Barrera, S. Gómez-Acevedo, M. García-Ortiz, y N. Gil-Luna, *Ficoweb PNALR: Base de Datos Orientada a la Catalogación y Divulgación de las Macroalgas Bénticas Marinas del Parque Nacional Archipiélago Los Roques Venezuela*, LXIV Convención Anual ASOVAC, Universidad Central de Venezuela, Facultad de Ciencias, Caracas, Venezuela, Noviembre 2014, https://www.researchgate.net/publication/271372854_ficoweb_pnalr
- [18] H. Niza, M. Bento, L. Mendes, A. Cartaxana, and A. Correia, *A Picture is Worth a Thousand Words: Using Digital Tools to Visualize Marine Invertebrate Diversity Data Along the Coasts of Mozambique and São Tomé & Príncipe*, Biodiversity Data Journal 9: e68817, May 2021, <https://doi.org/10.3897/BDJ.9.e68817>
- [19] I. Asaad, C. Lundquist, M. Erdmann, and M. Costello, *An Interactive Atlas for Marine Biodiversity Conservation in the Coral Triangle*, Earth System Science Data, vol. 11, no. 1, pp. 163-174, January 2019, <https://doi.org/10.5194/essd-11-163-2019>
- [20] GBIF.org, *Caribbean OBIS Node, Ocean Biodiversity Information System*, Administrative and technical contact Eduardo Klein, <https://www.gbif.org/publisher/1e9400fa-ce73-4737-9eaa-f25df148945c>
- [21] GBIF Secretariat & IAIA, *Buenas Prácticas para la Publicación de Datos sobre Biodiversidad Procedentes de Evaluaciones de Impacto Ambiental*, Colaboradores: Andrew Rodrigues, Dag Endresen, Rui Figueira, Cristina Villaverde, Miguel Vega, Nick King, Asha Rajvanshi, Jo Treweek, Copenhagen, February 2022, <https://doi.org/10.35035/doc-5xdm-8762>
- [22] S. Gómez-Acevedo y Y. Carballo-Barrera, *Benthic Macroalgae from the Coasts of Venezuela 1822-2021*, v1.12, Caribbean OBIS Node, Dataset/Occurrence, November 2022, <https://doi.org/10.15468/pjezmu>
- [23] GBIF.org, *Manual del Usuario del IPT*, Agosto 2024, <https://ipt.gbif.org/manual/es/ipt/latest/how-to-publish>
- [24] GBIF.org, *Estándares de Datos*, Agosto 2024, <https://www.gbif.org/es/standards>
- [25] BID-REX, *Mejores Datos, Mejores Decisiones: Cómo Aumentar el Impacto de la Información sobre Biodiversidad. Informe Técnico de la Fase 1 del Proyecto BID-REX - De los Datos sobre Biodiversidad a la Toma de Decisiones: Mejora del Valor Natural Gracias a Mejores Políticas de Desarrollo Regional*, Interreg Europe, España, ISBN: 978-84-09-10007-1, Marzo 2019.
- [26] Hugging Face, *Perplexity*, <https://www.perplexity.ai>
- [27] M. Troia and R. McManamay, *Filling in the Gaps: Evaluating Completeness and Coverage of Open - Access Biodiversity Databases in the United States*, Ecology and Evolution, vol. 6, no. 14, pp. 4654-4669, July 2016, <https://doi.org/10.1002/ece3.2225>
- [28] J. Gomeau, S. Kulkarni, F. Cala-Riquelme, and L. Esposito, *Measuring What We Don't Know: Biodiversity Catalogs Reveal Bias in Taxonomic Effort*, Bioscience, vol. 73, no. 2, pp. 112-123, February 2023, <https://doi.org/10.1093/biosci/biac116>
- [29] S. Gómez-Acevedo, Y. Carballo-Barrera, M. García-Ortiz, N. Gil-Luna, y A. Castillo, *Creación del Catálogo Taxonómico Nacional de las Macroalgas Bénticas Marinas de Venezuela*, Proyecto CDCH PG 03-8643-2013/1 2016, Diciembre 2016,

- https://www.researchgate.net/publication/312579335_Creacion_del_Catologo_Taxonomico_Nacional_de_las_Macroalgas_benticas_marinas_de_Venezuela
- [30] OData, *OData - The Best Way to REST*, <https://www.odata.org>
- [31] DCMI, *Dublin Core Metadata Initiative*, <https://www.dublincore.org/about/>
- [32] Schema.org, *Schema.org*, <https://schema.org>
- [33] TDWG, *Darwin Core*, Biodiversity Information Standards (TDWG), <https://dwc.tdwg.org>
- [34] D. De Pooter *et al.*, *Toward a New Data Standard for Combined Marine Biological and Environmental Datasets - Expanding OBIS Beyond Species Occurrences*. Biodiversity Data Journal 5., January 2017, <https://doi.org/10.3897/BDJ.5.e10989>
- [35] M. Jones, M. O'Brien, B. Mecum, C. Boettiger, M. Schildhauer, M. Maier, T. Whiteaker, S. Earl, and S. Chong. 2019. *Ecological Metadata Language (EML)*, version 2.2.0, KNB Data Repository, February 2022, <https://eml.ecoinformatics.org>
- [36] WoRMS, *World Register of Marine Species*, <https://www.marinespecies.org>
- [37] GBIF.org, *IPT: The Integrated Publishing Toolkit*, <https://www.gbif.org/ipt>
- [38] IODE, *Ocean Data Standards*, International Oceanographic Data and Information Exchange, <https://www.iode.org/index.php>
- [39] OpenArchives.org, *Open Archives Initiative. Protocol for Metadata Harvesting*, <https://www.openarchives.org/pmh>
- [40] Access to Biological Collection Data task group, *Access to Biological Collection Data (ABCD)*, Version 2.06, Biodiversity Information Standards (TDWG), 2007, <https://abcd.tdwg.org>
- [41] P. Yilmaz *et al.*, *Minimum Information About a Marker Gene Sequence (MIMARKS) and Minimum Information About Any (x) Sequence (MIxS) Specifications*. Nature biotechnology, vol. 29, no. 5, pp. 415-420, May 2011, <http://dx.doi.org/10.1038/nbt.1823>
- [42] J. Wiecek, D. Bloom, R. Guralnick, S. Blum, M. Döring, R. Giovanni, T. Robertson, and D. Vieglais, *Darwin Core: An Evolving Community-Developed Biodiversity Data Standard*. Plos One, vol. 7, no. 1, e29715, January 2012, <https://doi.org/10.1371/journal.pone.0029715>
- [43] OBIS.org, *Marine Biodiversity Data Mobilization Workshop 2024*, April 22-24 2024, <https://obis.org/2024/06/28/biodiv-mob-workshop-2024>
- [44] D. Hobern, A. Asase, Q. Groom, M. Luo, D. Paul, T. Robertson, P. Semal, B. Thiers, M. Woodburn, and E. Zschuschen, *Advancing the Catalogue of the World's Natural History Collections*, v2.0, Copenhagen, GBIF Secretariat, April 2020, <https://doi.org/10.35035/p93g-te47>
- [45] D. Hobern, L. Livermore, S. Vincent, T. Robertson, J.T. Miller, Q. Groom, and M. Grosjean, *Towards a Roadmap for Advancing the Catalogue of the World's Natural History Collections*, Research Ideas and Outcomes 8: e98593, December 2022, <https://doi.org/10.3897/rio.8.e98593>
- [46] SiB Colombia, *Actualización de las Plantillas del Estándar Darwin Core para Publicar Datos Sobre Biodiversidad*, Mayo 2024, <https://biodiversidad.co/post/2024/actualizacion-plantillas-dwc>
- [47] GoFair.org, *FAIR Guiding Principles for Scientific Data Management and Stewardship*, <https://www.go-fair.org/fair-principles>
- [48] OBIS.org, *The Obis Manual - 1.3 Data Policy*, <https://manual.obis.org/policy.html>
- [49] J. Lastimoso and W. Santiañez, *Updated Checklist of the Benthic Marine Macroalgae of the Philippines*. Philipp J Sci, vol. 150, no. S1, pp. 29-92, May 2021, <https://doi.org/10.56899/150.s1.04>
- [50] Y. Carballo-Barrera, M. García-Ortiz, S. Gómez-Acevedo, y N. Gil-Luna, *FICLAV: Una Clave Taxonómica para la División Ochrophyta (Algas Pardas) Utilizando Tecnologías Web y Árboles de Decisión*, XX Congreso Venezolano de Botánica, Universidad Nacional Experimental del Táchira, Táchira, Venezuela, Mayo 2013, https://www.researchgate.net/publication/271372749_ficlav_una_clave_taxonomica_para_la_division_ochrophyta_algas_pardas_utilizando_tecnologias_web_y_arboles_de_decision
- [51] S. Gómez, *Rhodophyta (Algas Marinas Rojas) del Parque Nacional Archipiélago Los Roques*, Tesis Doctoral, Facultad de Ciencias, UCV, Caracas, Venezuela, Noviembre 1998, https://www.researchgate.net/publication/383660557_Rhodophyta_algas_marinas_rojas_del_Parque_Nacional_Archipielago_Los_Roques_Venezuela
- [52] S. Gómez-Acevedo y M. García-Ortiz, *Macroalgas Bénticas Asociadas a Arrecifes Coralinos y Litorales Rocosos del Parque Nacional "Henri Pittier", Estado Aragua, Venezuela*, Bol. Inst. Oceanog. Venez., vol. 62, no. 01, pp. 17-36, September 2023, <http://dx.doi.org/10.13140/RG.2.2.18324.10881>
- [53] M. Guiry and G. Guiry, *AlgaeBase*, World-Wide Electronic Publication, University of Galway, August 2024, <https://www.algaebase.org>
- [54] N. González-Henríquez, *Los Bancos Genéticos y su Papel en la Conservación, Gestión y Trazabilidad de la Biodiversidad Marina en Iberoamérica: Caso Península de Santa Elena, Ecuador*, Foro Iberoamericano de los Recursos Marinos y la Acuicultura, no. 9, pp. 30-41, Abril 2020.
- [55] N. Arakaki, P. Carbajal, D. Marquez-Corigliano, S. Suárez, P. Gil-Kodaka, K. Pérez-Araneda, y F. Tellier, *Genética de Macroalgas en el Perú: Diagnóstico, Guía Metodológica y Casos de Estudio*. Inf. Inst. Mar Perú, vol. 48, no. 4, pp. 594-609, Octubre-Diciembre 2021, <https://repositorio.imarpe.gob.pe/bitstream/20.500.12958/3648/1/Informe%2048-4%20articulo8.pdf>
- [56] E. Zepeda, *Algas Marinas Bajo la Lupa de la Transcriptómica*, Abril 2021, <https://avanceperspectiva.cinvestav.mx/algas-marinas-bajo-la-lupa-de-la-transcriptomica>
- [57] GBIF.org, *Global Registry of Scientific Collections. A Worldwide Catalogue of Scientific Collections*, <https://scientific-collections.gbif.org>
- [58] ONU, *ODS*, <https://www.un.org/sustainabledevelopment/es/objetivos-de-desarrollo-sostenible>
- [59] M. Pinzón, *Desarrollo de los Módulos de Gestión de Datos, Generación de Estadísticas y Consultas del Catálogo Taxonómico Digital de Macroalgas Bénticas Marinas "Ficoflora Venezuela"*, Trabajo Especial de Grado, Escuela de Computación, Facultad de Ciencias, Universidad Central de Venezuela, Caracas, Venezuela, Octubre 2015.
- [60] Y. Carballo-Barrera, S. Gómez-Acevedo, M. García-Ortiz, y N. Gil-Luna, *Desarrollo de las Aplicaciones Informáticas del Catálogo Taxonómico Digital de Macroalgas Bénticas: Ficoflora Venezuela*, Jornadas de Investigación y Extensión de la Facultad de Ciencias, UCV, Mayo 2016, https://www.researchgate.net/publication/303644078_Desarrollo_de_las_aplicaciones_informaticas_del_Catologo_Taxonomico_Digital_de_Macroalgas_Benticas_Ficoflora_Venezuela
- [61] A. Peralta-Brichtova, J. Scott-Frías, C. Carmona-Suarez, C. Rodríguez, J. Perez-Benítez, A. Lopez-Ordaz, B. Marquez-Rojas, C. Lira, S. Gómez-Acevedo, Y. Carballo-Barrera, B. Rodríguez, F. Cavada-Blanco, J. Delgado, and E. Klein, *Literature-Based Occurrences Data of Marine Species in Venezuela*, Biodiversity Data Journal, vol. 11, e98213, February 2023, <https://doi.org/10.3897/BDJ.11.e98213>

Análisis de Sentimientos Migratorios en Publicaciones de Venezolanos en Medios Sociales

Livia Borjas¹, Mauricio Morales¹, Miguel Zamora¹

liborjas@ucab.edu.ve, mauricio1999morales@gmail.com, miguelzamora.23.97@gmail.com

¹ Escuela Ingeniería de Informática, Universidad Católica Andrés Bello, Guayana, Venezuela

Resumen: El presente trabajo tuvo como propósito realizar análisis de sentimientos y opinión migratoria en publicaciones de venezolanos en medios sociales, con la finalidad de apoyar el proceso de toma de decisiones en cuanto al fenómeno migratorio que sufre la nación venezolana. En primera instancia se llevó a cabo una revisión documental en aproximadamente ocho (08) artículos relacionados a los modelos de clasificación de textos, lo que permitió reconocer a los algoritmos de clasificación y el aprendizaje automatizado como los métodos más utilizados. En la etapa de desarrollo se definieron dos escalas de categorización, una que reflejó las opiniones de los usuarios y la otra los sentimientos asociados a su posición. Finalmente se procedió al proceso de construcción de los modelos de clasificación mediante la implementación de una aplicación móvil, donde los resultados obtenidos del análisis mostraron tendencias hacia la tristeza como sentimiento predominante y a la postura a favor de la migración, en cuanto a la opinión de los individuos.

Palabras Clave: Análisis de Sentimientos; Minería de Opinión; Clasificación de Textos; Algoritmos de Aprendizaje Automatizado; Fenómeno Migratorio.

1. INTRODUCCIÓN

El tema del análisis de sentimientos y minería de opinión ha tomado mucho interés en la actualidad. Muchas investigaciones han demostrado el creciente interés sobre el estudio de los fenómenos migratorios aplicando minería de datos con inteligencia artificial sobre publicaciones en redes sociales [1]. La presente contribución gira en torno al tema de analizar la opinión y sentimiento de los venezolanos expresados en sus redes sociales. El documento expone los primeros resultados de este estudio, organizando el aporte de la siguiente manera: (2) la sesión del “Planteamiento del Problema”, en la sección (3) se describen los “Antecedentes” de la investigación, en la sección (4) se presenta una propuesta de “Metodología” híbrida para aplicar en el caso en estudio, en la sección (5) se describen los resultados, mientras que en la sección (6) se presentan las conclusiones y recomendaciones, finalmente se presentan las “Referencias” en la sección (7).

2. PLANTEAMIENTO DEL PROBLEMA

El continuo deterioro de la economía venezolana a lo largo de los años, ha generado escasez severa de alimentos y medicamentos básicos que, sumada a la crisis política e institucional, resultaron en un movido flujo migratorio de venezolanos, reseñándose la diáspora de mayor volumen en la historia venezolana. Según Arias, Arias, Morffe, Martínez y Carreño, [2], en su artículo “Informe sobre la movilidad humana venezolana II”, mencionan que la migración internacional venezolana es uno de los fenómenos con mayor impacto en Latinoamérica, donde no sólo se ve afectada la población venezolana, sino también los países receptores, ya que, si la cifra de migración mantiene el ritmo actual, creará cada vez más problemas presupuestarios, decaimiento de servicios básicos y la atención sanitaria en dichos países.

En el artículo “Refugiados y Migrantes de Venezuela” de la Plataforma de Coordinación Interagencial para Refugiados y

Migrantes de Venezuela en su reporte de septiembre del 2021 [3] afirma que, la cifra de venezolanos que han emigrado es de aproximadamente 5,66 millones de individuos, que representan alrededor de un 20% de la población total venezolana.

Para llevar a cabo estudios sobre estos flujos migratorios, el Observatorio Proyecto Migración Venezuela, en su artículo “Percepción en redes sociales sobre la migración venezolana” [4] plantea que el análisis de las redes sociales es una de las herramientas más rápidas y eficaces para obtener una visión general de la situación, ya que proporciona información que posteriormente facilita a los encargados del proceso de toma de decisiones, la materialización de sus objetivos.

Según Hütt [5] en su artículo “Las redes sociales: Una nueva herramienta de difusión” el impacto de las redes sociales tiene la capacidad de cambiar la percepción que los usuarios tienen sobre un tema en particular, y en una situación como la crisis migratoria de Venezuela, las ideas y creencias formadas por la opinión pública pueden entorpecer la integración de los migrantes, por ello es importante aclarar las posturas actuales que se tienen sobre este fenómeno.

Para realizar el análisis de mensajes en redes sociales, una de las técnicas más utilizadas es el Análisis de Sentimientos también conocida como Minería de Opinión, que Liu [6] en su artículo “*Sentiment analysis and opinion mining*” define como “el campo de estudio que analiza las opiniones de las personas, sentimientos, evaluaciones, apreciaciones, actitudes y emociones sobre entidades tales como productos, servicios, organizaciones, individuos, cuestiones, eventos, tópicos y sus atributos”.

La relevancia del análisis de sentimientos radica en su capacidad para utilizar técnicas de procesamiento del lenguaje natural, sobre grandes volúmenes de información de forma automatizada, lo que facilita el reconocimiento de tendencias y

patrones para la categorización de las opiniones y los sentimientos de la población.

2.1 Propuesta

Dada la relevancia del tema migratorio venezolano unido con el aporte que puede sumar el análisis de sentimientos en redes sociales, el objetivo del presente trabajo se enfoca en realizar un proceso de análisis de sentimientos y minería de opinión respecto al tema de migración sobre publicaciones de venezolanos en medios sociales.

Para alcanzar este propósito se aplicaron los siguientes pasos: examinar algunos antecedentes relacionados con el análisis de sentimientos y minería de opinión en redes sociales a fin de establecer la manera efectiva de aplicar este procedimiento en el tema migratorio, determinar los criterios sobre los cuales se llevará a cabo dicho análisis, diseñar modelos descriptivos o predictivos de sentimientos y de opinión que se ajusten a los criterios establecidos, para finalmente implementar dichos modelos mediante una aplicación móvil para su publicación, explotación y uso.

3. ANTECEDENTES

El avance de las tecnologías de big data y minería de textos, en la última década, han demostrado los aportes que pueden obtenerse mediante la aplicación de análisis de sentimientos. A continuación, se describen algunas investigaciones que proporcionan base para el desarrollo del trabajo a realizar.

Al momento de realizar el análisis de sentimientos u opiniones, uno de los mayores retos es el de la comprensión de los tonos lingüísticos en un mensaje, en ese apartado Salaz [7], en su tesis doctoral titulada “Detección de patrones psicolingüísticos para el análisis de lenguaje subjetivo en español”, propuso un método para la detección de patrones psicolingüísticos en el análisis de sentimientos y la detección de la sátira en español.

Este método permite, a través de un enfoque automatizado supervisado, clasificar textos como positivo, negativo, neutro, muy positivo o muy negativo y como satíricos y no satíricos. Esta investigación demuestra las aplicaciones de los enfoques supervisados para reconocer aspectos particulares, en este caso el tono lingüístico en los mensajes.

Por su parte, Cestari [8] realizó una tesis de pregrado titulada “Propuesta para automatizar la asociación de emociones a textos en español”, la cual consistía en automatizar la asociación de frases en español a emociones, mediante el análisis de sentimientos. Tomando en cuenta el modelo de emociones de Ekman, se construyó un prototipo capaz de asociar las frases a un conjunto de seis (6) emociones, basado en algoritmos de aprendizaje de máquina, aprendizaje profundo y redes neuronales.

Este estudio demuestra la aplicación del análisis de sentimientos y algoritmos de aprendizaje de máquina a textos en español, así como la flexibilidad que ofrecen estos modelos ya que el número de categorías pueden variar según el propósito para el cual esté dirigido.

Arango y Osorio [9] realizaron un artículo de investigación titulado “Aislamiento social obligatorio: un análisis de sentimientos mediante machine learning”, donde se planteó

analizar los sentimientos subyacentes de los comentarios de Twitter relacionados con el aislamiento, identificando los temas y palabras frecuentemente utilizados en el contexto, a través de un algoritmo de machine learning. Se obtuvo como resultado la identificación del miedo como el sentimiento predominante durante todo el periodo de confinamiento.

Este artículo de investigación demuestra que el uso del análisis de sentimientos puede ir más allá que solo clasificar los mensajes en grupos particulares, sino que también permite identificar los temas y palabras características de cada uno de los grupos resultantes.

Abordando más afondo el fenómeno migratorio, se llevó a cabo una investigación sobre crisis migratoria basada en análisis de sentimientos [10] donde se plantea que la dinámica migratoria se ha producido a razón de crisis económicas, políticas y humanitarias. Por medio de un acercamiento desde el análisis de sentimientos se obtuvieron que las preocupaciones radicaban mayormente en las subcategorías de los derechos humanos, seguridad y desempleo. En base a estos resultados se puede validar que el análisis de sentimientos permite obtener opiniones públicas sobre un tema o situación determinada, para así tomar las medidas preventivas o correctivas correspondientes.

Adicionalmente, al aplicar la técnica de revisión documental, sobre aproximadamente más de ocho (08) referencias, [5][6][8][9][10][11][12][16], cuyas fechas de publicación estén vigentes con una década de diferencia al 2021 (fecha que se realizó la primera fase del estudio), en los cuales se aplica las técnicas de aprendizaje automático para realizar análisis de sentimientos. En este análisis documental se pudo apreciar que todos los trabajos abordaron un procedimiento metodológico basado en el entrenamiento de clasificadores de textos, aplicando las siguientes etapas;

- Etapa de recolección de datos
- Pre procesamiento de textos
- Entrenamiento de los modelos
- Despliegue de los resultados.

Para poder llevar a cabo la fase de recolección de datos, en su mayoría se basaron en mensajes de la red social Twitter, y para su extracción se utilizaron, el software estadístico R y en otras ocasiones lo hicieron a través de la API de Twitter con el lenguaje de Python.

De la misma forma para llevar a cabo la construcción de los clasificadores de textos se seleccionaron diferentes algoritmos basados en aprendizaje automático supervisado, entre los más utilizados se encontraron los siguientes algoritmos:

- Naive Bayes
- Máquina de Soporte Vectorial
- Regresión Logística
- Árboles de decisión
- K-vecinos.

Para evaluar el desempeño de los modelos construidos, se destacó el uso de las siguientes métricas de rendimiento:

- Precisión

- Recall
- fl-core.

En resumen, para llevar a cabo un estudio de análisis de sentimiento y opinión es importante tener presente:

- El uso de la plataforma de Twitter u otra red social como fuente de información, si es adecuado para el tema a indagar
- El uso del lenguaje Python o R para la extracción de textos de las redes sociales
- Las diferentes etapas que constituyen la construcción, entrenamiento y evaluación de modelos de clasificación de textos

El enfoque más utilizado es el de aprendizaje automático, más en concreto los basados en algoritmos de clasificación tradicionales.

4. METODOLOGÍA

El análisis de sentimientos y minería de opinión son temas que vienen en ascenso debido al impacto que pueden tener a nivel comercial o social. Debido a esto se decidió abordar la metodología CRISP-DM (Cross-Industry Standard Process for Data Mining o Proceso Estándar Para la Extracción de Datos en Todos los Sectores) [14][15], pero teniendo en consideración que el presente estudio tratará un tipo particular de minería de datos complejos, específicamente mensajes cortos de textos, se aplicará una propuesta de adaptación de autoría propia de la metodología CRISP-DM al caso de datos complejos de textos.

Esta adaptación consistirá de dos etapas generales: una etapa de pre-procesamiento y una etapa de descubrimiento. La primera etapa, está formada por la comprensión del problema, la creación del corpus y la transformación de los textos a algún tipo de representación estructurada o semiestructurada que facilite su posterior procesamiento, mientras que la segunda etapa, está constituida por el proceso de modelado para la clasificación de mensajes, la evaluación de rendimiento y finalmente el despliegue de los conocimientos obtenidos. En la Figura 1 se pueden apreciar las etapas para la metodología CRISP-DM adaptadas al proceso de minería de textos.

A continuación, se describen las etapas que se llevan a cabo, siguiendo la metodología CRISP-DM adaptada al proceso de análisis de sentimientos, como se puede observar en la Figura 1.

4.1 Comprensión del Negocio

En esta primera etapa, se lleva a cabo la etapa de revisión documental relacionada con el análisis de sentimientos, las técnicas de procesado de lenguaje y la construcción de modelos de clasificación, para fortalecer los conceptos fundamentales presentes en el análisis.

4.2 Definición de los Criterios para el Análisis de Sentimientos

En esta etapa se establecen los parámetros que se tendrán en cuenta para aplicar el análisis de sentimientos, así como las premisas que deben cumplir cada uno de ellos para la selección de sus elementos.

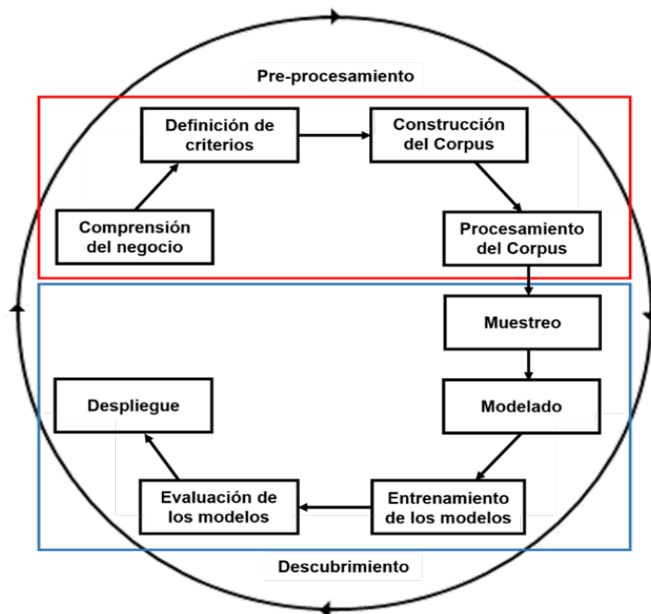


Figura 1: Adaptación de la Metodología CRISP-DM al Proceso de Análisis de Sentimientos

4.3 Construcción del Corpus

En esta etapa se recopilan la máxima cantidad de textos posibles que cumplan con la serie de criterios establecidos previamente y se asocia cada uno de ellos con alguna de las categorías de la escala de sentimientos.

4.4 Procesamiento del Corpus

En esta instancia se llevan a cabo un conjunto de técnicas para simplificar los mensajes a representaciones estructuradas, con la intención de mejorar la interpretación del análisis automatizado y así obtener mayor precisión al momento de entrenar y evaluar los modelos de clasificación.

4.5 Muestreo del Corpus

En esta etapa se define la proporción de los mensajes del corpus que formarán parte de los datos de entrenamiento, utilizados en la construcción del modelo, y los datos de prueba, utilizados en la etapa de evaluación del modelo.

4.6 Modelado y Entrenamiento de los Clasificadores

En esta etapa se lleva a cabo el desarrollo de los modelos, a partir de los algoritmos de clasificación seleccionados y la data de entrenamiento establecida.

4.7 Evaluación de los Modelos de Clasificación

En esta etapa se mide el desempeño de los modelos de clasificación haciendo uso de los datos de prueba. Este proceso se logra mediante el estudio de las métricas de rendimiento que evalúan el índice de éxito del modelo para diferentes situaciones.

4.8 Despliegue de los Modelos de Clasificación

Finalmente, según los resultados obtenidos, los modelos pueden ser desplegados para su visualización o utilizados como base para otros trabajos con relación al área.

5. RESULTADOS

Los resultados de los objetivos más resaltantes, después de aplicar la metodología, se describen a continuación:

5.1 Definición de Criterios para el Análisis de Sentimientos

Entre los criterios a tomar en cuenta para el proceso de recolección de textos, que constituye la primera fase para llevar a cabo un análisis de sentimientos, se encuentran:

a) *Selección de la Plataforma.* La plataforma seleccionada para la extracción de los textos de estudio fue Twitter debido su facilidad de acceso a la API para el momento del estudio, año 2021, los formatos al realizar una extracción de datos y como indica Mathews et al. (2016) citado por Saura, Palos y Ríos [16], la red social Twitter, además de generar gran cantidad de datos y mensajes, permiten vincularlos a un evento o característica y posteriormente segmentar audiencias al recopilar mensajes en torno a un hashtag, las cuales son etiquetas ubicadas dentro de los textos para asociarlos a algún tema en particular.

b) *Selección de la Población Objetivo.* En esta etapa se seleccionaron los perfiles de twitter, de los cuales se extraerán los mensajes para el análisis. Para definir a la población se utilizaron los criterios mencionados por Arias, Villasis y Miranda [1] donde establecen que para definir una población es necesario tener en cuenta las características de homogeneidad, la temporalidad y los límites espaciales. Teniendo en cuenta estos factores, para la selección de la población objetivo se definieron las siguientes premisas:

- Homogeneidad. Los perfiles deben pertenecer a individuos mayores a 18 años de edad. Los perfiles deben tener alguna participación sobre la migración
- Temporalidad. Los perfiles deben haber tenido participación en el año 2021, año donde se realiza la recolección de los datos a través del API
- Definir los límites espaciales. Los perfiles deben pertenecer a individuos venezolanos

Para llevar a cabo esta tarea se utilizó la herramienta web FollowerSearch, la cual es una herramienta de análisis de twitter, especializada en la búsqueda de perfiles al suministrar ciertos parámetros. Para este estudio, el criterio de selección aleatoria, de perfiles de venezolanos y mayores de edad. Haciendo uso de ella se obtuvo un total de 128 usuarios de twitter que cumplieran con las características requeridas.

c) *Definición de la Escala de Opiniones y Sentimientos.* Para obtener la clasificación más precisa sobre los mensajes a analizar se determinó necesario el uso de dos escalas de categorización. La primera se basa en una escala de Likert, que según Matas [17], es uno de los instrumentos más utilizados para la medición de opiniones, en este caso se determinó el uso de tres ítems, ya que este plantea que no se encuentran diferencias significativas en las versiones de tres, cinco y siete alternativas, y para esta temática no resulta relevante la subclasificación de las posturas:

- A favor de la migración.
- Neutral ante la migración.

- En contra de la migración.

La segunda escala hace referencia al sentimiento asociado a la opinión emitida, en esta investigación se tomó en cuenta la clasificación propuesta por Jack, Garrod y Schyns [18] de la Universidad de Glasgow, donde sugieren que las categorías utilizadas para medir las emociones estarán formadas por las siguientes categorías:

- Felicidad
- Tristeza
- Ira
- Miedo.

d) *Enfoque de análisis.* En el momento del análisis, los métodos de aprendizaje automático tenían mayor relevancia debido a que utilizan las capacidades de un algoritmo de clasificación para que encuentre por sí mismo a qué polaridad pertenecen las palabras y oraciones, a partir de un conjunto de datos de entrenamiento, debido a esto, la investigación se centrará en este tipo de enfoque.

e) Algoritmos de Aprendizaje Automático

Con el fin de aplicar los modelos de aprendizaje automático es necesario realizar la selección de los algoritmos de clasificación que se tomarán en cuenta al momento de construir el modelo ya que cada uno de los algoritmos de clasificación tiene una forma distintiva para llevar a cabo la categorización de los datos. En la Tabla I, se pueden observar los algoritmos seleccionados y una breve descripción de cada uno de ellos.

Tabla I: Algoritmos de Clasificación para el Análisis de Sentimientos

Algoritmo	Descripción
Naive Bayes (NB)	Es un modelo basado en la identificación de características para cada una de las categorías. Se basa en la predominancia de un grupo de características al realizar una predicción.
Maximum Entropy (MaxE)	Es un modelo basado en regresión logística, utiliza el principio de la máxima entropía, para determinar las características mutuamente dependientes que identifican una categoría.
Árboles de Decisión (DT)	Es un modelo basado en el uso de una serie de reglas simples extraídas de los datos para obtener así una predicción verificando el cumplimiento de alguna de las obtenidas.
Máquina de Vectores de Soporte (SVM)	Es un modelo basado en la identificación de planos que le permitan delimitar las diferentes categorías

5.2 Construcción del Corpus

Para recopilar los mensajes de la plataforma de Twitter, se utilizó la herramienta Twint, que es una librería del lenguaje Python para el scraping de tweets, la cual tuvo como resultado 5000 tweets, donde se incluían los más recientes de cada uno

de los individuos de la población objetivo. Estos mensajes fueron almacenados posteriormente en un archivo .CSV, debido a la facilidad para operar sobre él con las herramientas disponibles en el lenguaje Python.

Posteriormente se llevó a cabo la limpieza del corpus que engloba múltiples etapas, la primera de ellas, es la selección de las columnas que serán utilizadas en el análisis las cuales son las siguientes:

- Fecha de publicación
- Usuario que lo publicó
- Contenido del tweet

Luego se procedió a la eliminación de aquellos tweets que no aportan valor al desarrollo de este trabajo como son los siguientes casos

- Tweets repetidos
- Tweets que no tienen como enfoque la migración venezolana

Como resultado se descartaron un total de 1254 mensajes, resultando una muestra de 3746 tweets que conformaron el corpus inicial. Finalmente, para concluir esta etapa, se procedió a categorizar cada uno de los mensajes, en las escalas de sentimientos anteriormente descritas. En la Tabla II y Tabla III se muestran la distribución de los tweets para cada una de las clasificaciones.

Tabla II: Distribución de los Mensajes Según su Opinión sobre la Migración

Postura sobre la migración	Número de tweets
A favor de la Migración	2702
Neutral ante la Migración	191
En contra de la Migración	853

Tabla III: Distribución de los Mensajes según la Emoción Asociada

Emociones	Número de tweets
Felicidad	871
Tristeza	913
Ira	878
Miedo	1084

5.3 Pre-procesamiento del Corpus

Después de obtener el corpus inicial se llevó a cabo el procesamiento del lenguaje natural, aplicando técnicas de análisis sobre los textos de entrada que permiten mejorar la capacidad de interpretación de los algoritmos, en este estudio se llevaron a cabo las siguientes etapas:

a) Tokenización

Se utilizó el paquete de nltk.tokenize para separar cada una de las palabras y signos de puntuación que conforman un mensaje y así analizar cada una de estas por separado en las siguientes etapas.

b) Eliminación de Ruido

En esta etapa haciendo uso del paquete nltk.corpus, se removieron los siguientes casos que afectaban el análisis

- Palabras de Parada
- Espacios múltiples
- Enlaces web
- Menciones a otros usuarios
- Emoticones

c) Normalización

En esta etapa haciendo uso del paquete nltk.stem se aplicaron las siguientes técnicas de normalización

- Conversión de todos los tokens extraídos a letras minúsculas
- Stemming, donde cada una de las palabras es transformada a su raíz.

5.4 Muestreo del Corpus

Con la finalidad de evaluar cada uno de los modelos de clasificación seleccionados se realizaron tres pruebas, con base en el corpus creado, las cuales fueron variadas en el porcentaje de mensajes utilizados para su entrenamiento. En la Tabla IV, se pueden apreciar las pruebas realizadas y la distribución de los mensajes utilizados en cada apartado.

Tabla IV: Distribución del Corpus para las Pruebas de los Modelos

Nº Prueba	Entrenamiento (%)	Entrenamiento (Totales)	Prueba (%)	Prueba (Totales)
1	70	2622	10	375
2	80	2997		
3	90	3371		

Los mensajes que forman parte de la categoría de entrenamiento, son aquellos a partir de los cuales se creará el modelo de clasificación, mientras los que pertenecen a la categoría de prueba son aquellos utilizados para evaluar el desempeño de los modelos posteriormente.

5.5 Entrenamiento de los Modelos de Análisis de Sentimientos

Una vez construido el corpus y desarrollados los algoritmos de clasificación, se procedió a realizar la etapa de entrenamiento del algoritmo, en la cual se les suministraron a los algoritmos los datos ya procesados y etiquetados, para así obtener los modelos clasificadores que serán los encargados posteriormente de realizar las predicciones y categorizar nuevos elementos.

5.6 Evaluación de los Modelos de Análisis de Sentimientos

En esta etapa se llevaron a cabo la serie de pruebas descritas en la sección del muestreo del corpus, que permitieron estimar el desempeño de los modelos de clasificación, las cuales variaron el porcentaje de mensajes utilizados para su entrenamiento. Al evaluar los resultados de las métricas de rendimiento descritas en la Tabla V.

Tabla V: Métricas de Rendimiento para Evaluar los Modelos de Clasificación

Métrica	Descripción
Precisión	Representa el porcentaje de mensajes predichos de una categoría que pertenecen a ella realmente
Sensibilidad	Representa el porcentaje de mensajes reales de una categoría que fueron correctamente asignados en la predicción
Valor F1	Es un valor que representa la precisión y la sensibilidad en un único atributo
Exactitud	Representa el porcentaje de casos que el modelo ha acertado de todas las categorías

En la Tabla VI, se muestra el resumen de los resultados obtenidos en la evaluación de rendimiento para cada uno de los modelos de clasificación.

Tabla VI: Rendimiento de los Modelos de Clasificación en la 3era Prueba

Nº Prueba	Precisión (%)	Sensibilidad (%)	Valor-F1 (%)	Exactitud (%)
Naive Bayes	79.23	76.56	77.27	77.07
Maximum Entropy	80.27	79.5	79.76	79.73
Árboles de Decisión	76.96	76.02	76.35	76.27
Máquinas de Vectores de Soporte	83.08	82.74	82.77	82.67

El rendimiento de los modelos de clasificación fue mejorando con el incremento de la data de entrenamiento, lo cual se cumplió para cada uno de los algoritmos a lo largo de las tres pruebas realizadas. Por ello, para el análisis de resultados se tomó como referencia principal la última prueba que utilizó la mayoría de los datos de entrenamiento.

El modelo que obtuvo mejores resultados bajo estas condiciones, fue el basado en máquinas de vector de soporte, ver Tabla VI, que obtuvo una tasa de éxito en el 82.67% de los casos, que para la prueba realizada representó un total de 310 predicciones correctas de 375 textos seleccionados. Debido a esto, el modelo de clasificación que se utilizó principalmente en la aplicación desarrollada fue el modelo basado en máquinas de vector de soporte. Sin embargo, también se contempló la posibilidad de utilizar los otros modelos para observar los distintos comportamientos.

Tabla VI: Resultados para la Escala de Sentimientos con la Data de Entrenamiento al 90%

Categoría	precision	recall	f1-score	confusion matrix
Felicidad	85	78.16	81.44	68 2 5 12
Ira	89.89	90.91	90.4	2 80 1 5

Miedo	87.25	81.65	84.36	2 4 89 14
Tristeza	70.19	80.22	74.87	8 3 7 73
average	83.08	82.74	82.77	accuracy 82.67

Desde el punto de vista de la escala de sentimientos y opiniones, se puede determinar el rendimiento general que tuvieron cada una de las categorías a través de los múltiples modelos de clasificación, y así reconocer las tendencias generales de los modelos. En la Figura 2 podemos apreciar que la opinión a favor de la migración fue la que tuvo el mejor desempeño, mientras que las otras posturas no lograron un buen rendimiento, esto viene dado principalmente a la diferencia en la cantidad de mensajes de entrenamiento disponibles en el corpus inicial, donde hay una predominancia de textos asociados a la categoría a favor de la migración lo que permite un modelo más refinado para detectar este tipo de textos y genera una predisposición a su clasificación.

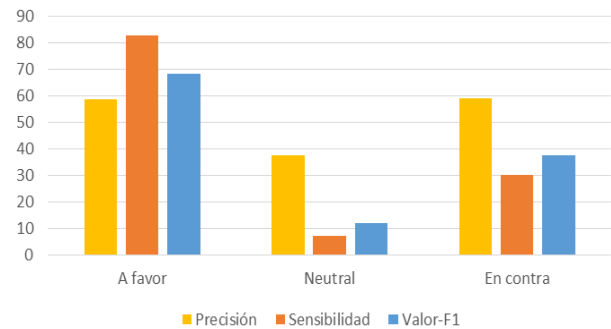


Figura 2: Rendimiento General según la Escala de Opiniones

En la Figura 3 por otro lado, podemos observar que las categorías de la escala de sentimientos se encuentran todas con un desempeño similar, esto se debe principalmente, a que la cantidad de datos de entrenamiento suministrados fueron similares para cada uno de los sentimientos.

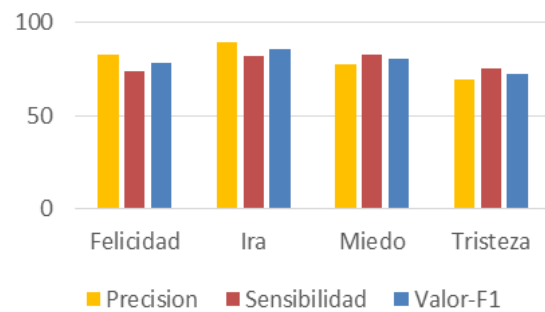


Figura 3: Rendimiento General según la Escala de Sentimientos

5.7 Resultados del Análisis de Sentimientos

Para obtener una visión general de la situación actual, se decidió analizar la mayor cantidad de registros de data, en vista que mejoraba el comportamiento del modelo, por lo cual se utilizó un total de 17.252 mensajes disponibles referentes al fenómeno de migración. Haciendo uso del modelo de clasificación basado en máquinas de soporte vectorial, que fue

el que obtuvo el mejor de desempeño de los modelos construidos. Los resultados fueron los siguientes:

a) Distribución de Sentimientos General

Con la finalidad de conocer la predominancia de alguna de las categorías de la escala de opiniones y sentimientos, en las Figuras 4 y Figura 5 se disponen la distribución de los resultados obtenidos en la clasificación.

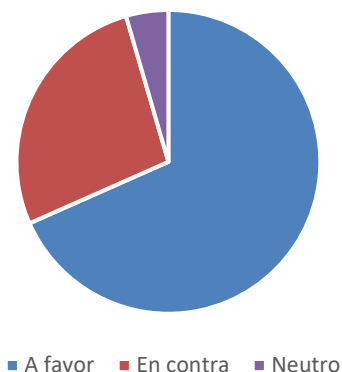


Figura 4: Distribución de la Escala de Opiniones

En la Figura 4 se puede observar que existe una tendencia de opinión clara hacia la postura a favor de la migración, el modelo clasificó aproximadamente un 64.94% de los mensajes a esta categoría, mientras que un 25.77% fueron asignados a la postura en contra y tan solo un 4.28% a la postura neutral.

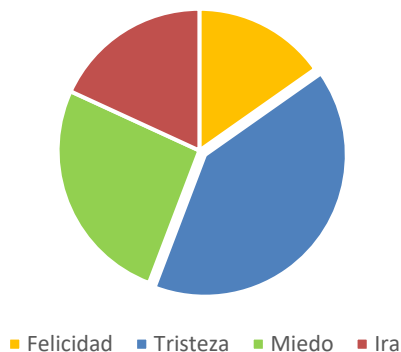


Figura 5: Distribución de la Escala de Sentimientos

Por otro lado, en cuanto a la escala de sentimientos, en la Figura 5 se puede observar una distribución más equilibrada, que se decanta ligeramente hacia la emoción de tristeza con un total de 40.55% de los mensajes, mientras que la distribución de los mensajes asignados a las categorías de miedo, ira y felicidad fueron similares con un total de 26.07%, 18.14% y 15.14% de los textos respectivamente.

b) Temas y Palabras Frecuentes

Los resultados de este estudio nos brindaron una mayor perspectiva sobre cuáles son los pensamientos de la población venezolana teniendo en cuenta su postura, para su obtención se evaluaron dos tipos de variables, los Unigramas y los N-gramas. Los Unigramas, en este caso, están representados por palabras, por ende, este método nos permite conocer aquellos

términos que tienen mayor aparición en los textos, en la Figura 6 se pueden observar una lista con los Unigramas más relevantes y su número de apariciones. Por otro lado, los N-gramas están representados en este ámbito por un conjunto de palabras, que proporcionaron una perspectiva más detallada, en la Figura 7 se pueden observar los N-gramas que mayor aparición tuvieron y su número de apariciones.

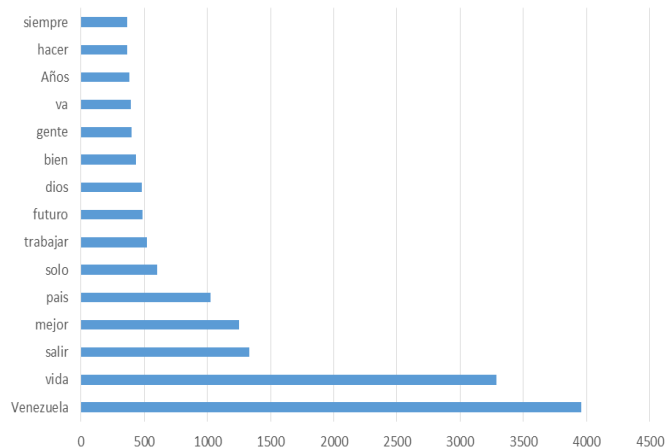


Figura 6: Unigramas con Mayor Aparición en el Análisis de Sentimientos General

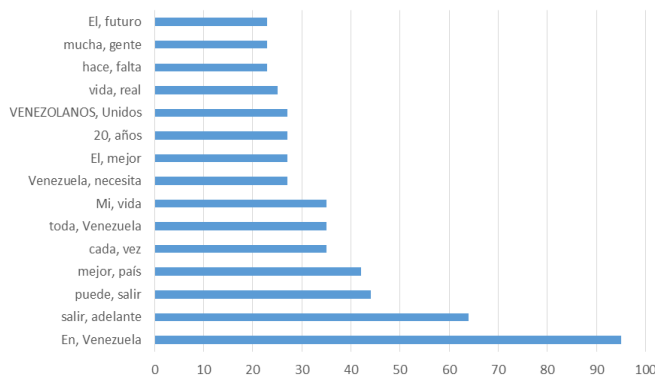


Figura 7: N-gramas con Mayor Aparición en el Análisis de Sentimientos General

Para identificar las relaciones que tienen estos términos con cada una de las categorías de la escala de opiniones, en las Figura 8, Figura 9 y Figura 10 se presentan los N-gramas de las posturas a favor, en contra y neutral ante la migración respectivamente, con la intención de comparar los valores característicos de cada uno.

Se puede observar que los términos más relevantes de la postura a favor de la migración se enfocan a mencionar Venezuela, la vida que estos tengan en el país y la posibilidad de salir del mismo. En cuanto a la postura en contra de la migración, esta presenta un enfoque más centrado en la mejora de Venezuela y la contribución del pueblo venezolano para lograrlo. Y finalmente la postura neutral ante la migración solo hace presencia en un par de palabras debido a que el mayor número de los mensajes se ven identificados en las 2 posturas anteriores.

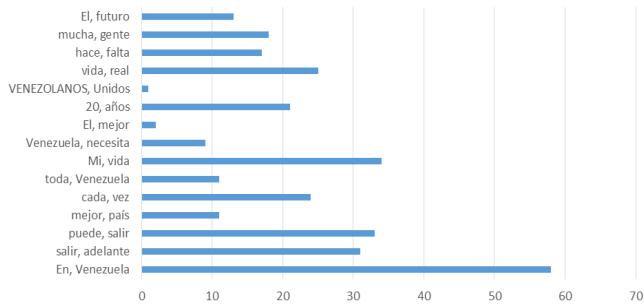


Figura 8: N-gramas con Mayor Aparición en Textos a Favor de la Migración

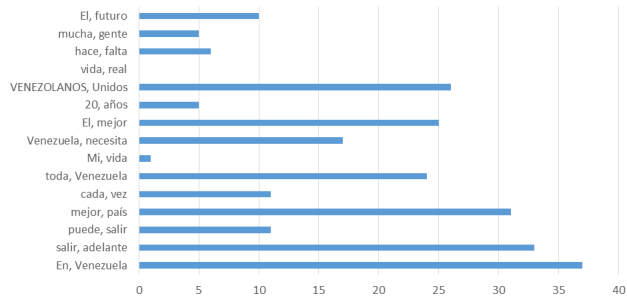


Figura 9: N-gramas con Mayor Aparición en Textos en Contra de la Migración

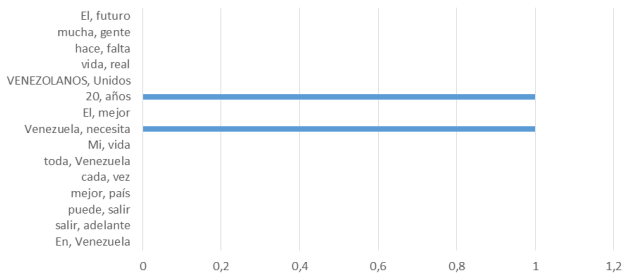


Figura 10: N-gramas con Mayor Aparición en Textos Neutrales ante la Migración

5.8 Desarrollo de la Aplicación

Para el desarrollo del aplicativo se llevó a cabo una serie de procesos, entre ellos se destacan, la creación de la base de datos, que fue la estructura encargada de almacenar los datos necesarios para la generación y evaluación de los modelos de clasificación, la construcción de una API, que fue la encargada de conectar la aplicación móvil con el servicio de base de datos y finalmente la aplicación que utilizó ambas herramientas para mostrar los resultados del análisis obtenido.

a) Base de Datos

En primera instancia, se diseñó una base de datos capaz de almacenar los valores necesarios para la etapa de entrenamiento y evaluación de los modelos de clasificación, debido a la simple naturaleza de los datos utilizados, el manejador de base de datos seleccionado fue PostgreSQL, una base de datos de tipo relacional, la cual provee un gran manejo y facilidad de acceso desde Python.

b) Application Programming Interface (API)

Para la comunicación entre la aplicación móvil y el servidor de base de datos, se desarrolló un servicio web que serviría de

intermediario en este proceso. A través de la Tabla VII, se describen las rutas para atender a las peticiones de la aplicación.

Tabla VII: Rutas del Servicio Web Intermediario entre la Aplicación y el Servidor de Base de Datos

Ruta	Método HTTP	Descripción
/tweets	GET	Petición que devuelve todos los tweets de data de entrenamiento o los ya evaluados que estén almacenados en la Base de datos.
/tweets	POST	Petición que permite añadir un registro en la Base de Datos.
/tweets/scraping	GET	Petición que ejecuta la función para poder recolectar los datos de twitter y los procesa para ser almacenados en un archivo csv.
/tweets/classification	GET	Petición que ejecuta los modelos de predicción usando los datos de ejemplo para demostrar el funcionamiento de los mismos.
/tweets/classification/message	POST	Petición que recibe un mensaje y el nombre del usuario, para poder ejecutar los modelos predictivos sobre ese mensaje.
/tweets/fill	GET	Petición que solicita que se vacíen y rellenen la tabla de la base de datos.
/tweets/ngrams	GET	Petición para recolectar la frecuencia de las palabras que estén en los datos de prueba.
/statistics /general_probability_distribution	GET	Petición que devuelve el porcentaje de cuantos mensajes fueron clasificados en cada una de las escalas de sentimientos y los de opinión.
/statistics /distribution_by_date	GET	Petición que devuelve la cantidad de ocurrencias de cada una de las escalas de sentimiento y los de opinión pero esta vez bajo el parámetro de los meses del año.

/statistics /distribution_monograms	GET	Petición que devuelve las 15 palabras más frecuentes con sus datos de sentimientos y opinión según en los mensajes que haya aparecido.
/statistics /distribution_ngrams	GET	Petición que devuelve los 15 ngramas (Bigramas o Trigramas) más frecuentes con sus datos de sentimientos y opinión según en los mensajes que haya aparecido.

c) Aplicación Móvil

Para el desarrollo de la aplicación móvil se utilizó la herramienta de Flutter ya que como expone Herazo [19] es un marco moderno y sencillo para crear aplicaciones móviles, que tiene su propio motor de renderizado y permite ver los resultados en tiempo real.

El desarrollo de la aplicación se dividió en dos secciones principales, el desarrollo de la sección de las estadísticas del análisis de sentimientos y opiniones y la sección de predicciones utilizando los modelos de clasificación desarrollados. Para ambas secciones se hace uso de la API descrita en la Tabla VII que permite la interacción con el servidor de base de datos.

Para la representación de las gráficas, se utilizó la librería *charts_flutter*, la cual cuenta con un conjunto extensivo de gráficas para datos estadísticos.

6. CONCLUSIONES Y RECOMENDACIONES

A partir del análisis del problema en perspectiva de los resultados obtenidos en el trabajo, se extrajeron las siguientes conclusiones:

- La revisión del estado del arte indica que, para llevar a cabo el proceso de análisis de sentimientos y minería de opinión, la plataforma Twitter constituye una gran fuente de data, pues es de las redes sociales más utilizadas para la publicación de textos, los cuales pueden ser extraídos y procesados mediante las herramientas que brinda el lenguaje de Python, para la construcción del corpus inicial y posteriormente los modelos de clasificación.
- En la determinación de los criterios para el análisis de sentimientos y minería de opinión se identifican los parámetros que apoyan al proceso de recolección y análisis de los textos, entre ellos se destacan las características de la población objetivo, la escala de los sentimientos y opiniones que categorizan los textos y las métricas de rendimiento para evaluar el desempeño de los modelos que clasifican los textos relacionados al tema de la migración venezolana.
- Para diseñar modelos predictivos en base a textos publicados en redes sociales, es necesario la construcción de un corpus de palabras presentes en las publicaciones de

venezolanos para luego determinar los sentimientos y opiniones referentes al tema de la migración con el uso de las técnicas de procesamiento del lenguaje natural y los algoritmos de aprendizaje automatizado. La calidad de los datos recolectados afecta de manera proporcional la calidad y validez de los modelos extraídos.

- Esta investigación muestra el ventajoso uso de los algoritmos de clasificación y la selección de categorías para la escala de sentimientos y opiniones para la construcción de modelos de clasificación que abarquen todas las perspectivas de la población venezolana sobre el fenómeno migratorio. Una vez entrenado el clasificador, se lleva a cabo el proceso de evaluación, haciendo uso de las métricas de rendimiento que permiten determinar su desempeño al categorizar textos relacionados a la migración venezolana.
- El resultado de la evaluación de los modelos de clasificación muestra que el basado en el algoritmo de Máquina de Soporte Vectorial es el que posee el mejor desempeño, obteniendo una exactitud general del 82.67% al momento de clasificar los mensajes de venezolanos sobre el fenómeno migratorio. El uso de este modelo para realizar el análisis de sentimientos y minería de opinión sobre la población venezolana en el año 2021 respecto al fenómeno migratorio venezolano, destaca que el sentimiento predominante es la tristeza, con un porcentaje de aparición en textos del 40.55%, y en cuanto a las opiniones, la postura a favor de la migración prevalece en el 64.94% de los casos.
- La implementación de los diferentes modelos predictivos mediante una aplicación móvil resulta ser útil, ya que permite la publicación de los resultados del análisis de datos, como la distribución general de la escala de sentimientos y opiniones, su distribución a través del tiempo y las palabras clave que identifican a cada una de las categorías, mediante una plataforma potencial de uso masivo, que apoya el proceso de toma de decisiones, suministrando información oportuna, para afrontar el fenómeno migratorio.
- El resultado obtenido en el presente trabajo evidencia la factibilidad del uso del análisis de sentimiento y minería de opinión de publicaciones en redes sociales, para la construcción de un modelo predictivo sobre el fenómeno migratorio venezolano, demostrando la razón del aumento de este tipo de investigaciones para el descubrimiento y explotación de opiniones y sentimientos de una comunidad respecto a un evento en un momento determinado.
- En base a los resultados obtenidos en la presente investigación, a continuación, se enumeran una serie de recomendaciones a tomar en cuenta para trabajos futuros:
- Mejorar los modelos de clasificación, explorando la aplicación y/o combinación de otras técnicas de análisis de datos, como otras técnicas de agrupamiento, reglas de asociación, entre otras.
- Refinar el comportamiento de los modelos de clasificación descubiertos, utilizando un conjunto de datos

más elevado y equitativamente distribuido entre todas las categorías.

- Explorar el uso de otros medios sociales para la recolección de datos, ya que pueden tener una representación distinta de los diferentes grupos de la sociedad venezolana.
- Experimentar el uso de otras arquitecturas, modelos de entrenamiento, librerías y entornos de desarrollo, con la finalidad de ampliar el conocimiento sobre otras opciones de análisis de sentimientos, y así elegir la más adecuada para un problema en específico.
- Ampliar el alcance de la aplicación móvil implementada, para abarcar la interacción entre usuarios, a modo de red social, permitiendo una mejora en la exposición de información para la toma de decisiones.

REFERENCIAS

- [1] J. Arias, M. Villasis, y M. Miranda. *El Protocolo de Investigación III: La Población de Estudio*. Revista Alergia, México. 2016
- [2] R. Arias, N. Arias, M. Morffe, C. Martínez, y T. Carreño. *Informe sobre la Movilidad Humana Venezolana II*. 2019
- [3] Plataforma de Coordinación Interagencial para Refugiados y Migrantes de Venezuela. *Refugiados y Migrantes de Venezuela*. 2021
- [4] Observatorio Proyecto Migración Venezuela. *Percepción en Redes Sociales sobre la Migración desde Venezuela*. 2020
- [5] H. Hütt. *Las Redes Sociales: Una Nueva Herramienta de Difusión*. 2012. <https://www.redalyc.org/pdf/729/72923962008.pdf>
- [6] B. Liu. *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publishers. Chicago, United States. 2012
- [7] M. Salaz. *Detección de Patrones Psicolingüísticos para el Análisis de Lenguaje Subjetivo en Español*. 2017. https://rua.ua.es/dspace/bitstream/10045/74617/1/PLN_60_10.pdf
- [8] A. Cestari. *Propuesta para Automatizar la Asociación de Emociones a Textos en Español*. Trabajo de Grado. UCAB, Puerto Ordaz. 2019.
- [9] C. Arango y C. Osorio. *Aislamiento Social Obligatorio: Un Análisis de Sentimientos Mediante Machine Learning*. 2020. <http://revistasumadenegocios.konradlorenz.edu.co/vol12-num-26-2021-aislamiento-social-obligatorio-un-analisis-de-sentimientos-mediante-machine-learning/>
- [10] M. Arguedas, J. Beita, F. Rodríguez, J. Umaña y M. Vaca. *Crisis Migratoria en Colombia y Costa Rica: Una Visión desde el Análisis de Sentimientos*. 2020. <https://revistas.ucr.ac.cr/index.php/humanidades/article/view/42238>
- [11] Gonzáles. *Clasificador de Texto Mediante Técnicas de Aprendizaje Automático*. 2020
- [12] A. Mundalik. *Aspect Based Sentiment Analysis Using Data Mining Techniques Within Irish Airline Industry*. 2018. <http://norma.ncirl.ie/3413/1/aishwaryamundalik.pdf>
- [13] Páez y Monroy. *Implementación de un Modelo de Análisis de Sentimientos con Respecto a la JEP Basado en Minería de Datos en Twitter*. 2020
- [14] P. Chapman, J. Clinton, R. Keber, T. Khabaza, T. Reinartz, C. Shearer, R. Wirth. 2000. CRISP-DM 1.0 Step by step BGuide. Edited by SPSS.
- [15] F. Peralta. *Proceso de Conceptualización del Entendimiento del Negocio para Proyectos de Explotación de Información*. Revista Latinoamericana de Ingeniería de Software, 2(5): 273-306, ISSN 2314-2642. 2014
- [16] J. Saura, P. Palos-Sánchez y M. Ríos. *Un Análisis de Sentimiento en Twitter con Machine Learning: Identificando el Sentimiento sobre las Ofertas de #BlackFriday*. Revista ESPACIOS. Vol. 39 (Nº 42) Año 2018. Pág. 16. 2018.
- [17] A. Matas. *Diseño del Formato de Escalas tipo Likert: un Estado de la Cuestión*. 2018. http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1607-40412018000100038
- [18] R. Jack, O. Garrod, P. Schyns. *Dynamic Facial Expressions of Emotion Transmit an Evolving Hierarchy of Signals over Time*. 2014. <https://www.sciencedirect.com/science/article/pii/S0960982213015194>
- [19] L. Herazo. *¿Qué es Flutter y por qué Utilizarlo en la Creación de Apps Móviles?*. 2021. <https://anincubator.com/que-es-flutter-y-por-que-utilizarlo-en-la-creacion-de-apps-moviles/>



Índice de Autores

Índice de Autores

A

Arraiz Aleidys	10
Arroyo Marcelita	2

B

Barrios Judith	17
Borjas Livia	44

C

Carballo Yusneyi	10, 25, 32
------------------	------------

F

Fernández José	25
----------------	----

G

García Jesús	2
Gil Ángel	2
Gómez Santiago	32

M

Montilva Jonás	17
Morales Mauricio	44

T

Torrealba Miguel	6
------------------	---

Z

Zamora Miguel	44
---------------	----



CoNCISa 2024

Sociedad Venezolana de Computación

Memorias de la Décima Conferencia Nacional
de Computación, Informática y Sistemas



<https://concisa.net.ve/2024>