

International Journal of Information Science  
and Technological Applications-UAS

# IJISTA



Diciembre 2025 - Mayo 2026. Vol. II Núm. I

U N I V E R S I D A D A U T Ó N O M A D E S I N A L O A

ISSN (en trámite)

**UNIVERSIDAD AUTÓNOMA DE SINALOA**

**International Journal of Information Science and  
Technological Applications-UAS**

**IJISTA**

**Vol. II. Núm. I, Mayo 2026. ISSN(en trámite)**

**REVISTA**

**Facultad de Informática Culiacán**



**Culiacán de Rosales, Sinaloa, México.**

**<https://revistas.uas.edu.mx/index.php/IJISTA/index>**

# **DIRECTORIO INSTITUCIONAL**

**Dr. Jesús Madueña Molina**

Rector

**Dra. Nidia Yuniba Brunn Corona**

Secretario General

**Dr. Alfonso Mercado Gómez**

Director de Servicios Escolares

**Dr. Sergio Mario Arredondo Salas**

Secretario Académico Universitario

**Dr. Wenseslao Plata Rocha**

Vicerrector de Unidad Regional Centro

**Dra. Marcela de Jesús Vergara Jiménez**

Director General Investigación y Posgrado

**Dr. Joel Cuadras Urías**

Director General del Sistema Bibliotecario

**Dr. Roberto Bernal Guadiana**

Director de la Facultad de Informática Culiacán

**Dra. Xiomara Penélope Zaldívar Colado**

Editora en Jefe de la Revista IJISTA-UAS

## **Comité Editorial**

Dr. Roberto Bernal Guadiana  
roberto.bernal@uas.edu.mx  
**Director de la Revista**

Dra. Xiomara Penélope Zaldívar Colado  
editor.ijista@uas.edu.mx  
**Editor en Jefe**

MC. Raúl Quevedo García  
admi.ijista@uas.edu.mx  
**Administrador Open Journal System**

Dr. José de Jesús Uriarte Adrian  
gestor.ijista@uas.edu.mx  
**Gestor Editorial**

MC. Thania Roxaana Félix González  
difusion.ijista@uas.edu.mx  
**Editor de Difusión y Comunicación**

Dra. Cynthia Itzel Jiménez Bernal  
cynthiajimenez@uas.edu.mx  
**Corrector de Estilo**

Lic. Vladimir Nieves Cázarez  
traductor.ijista@info.uas.edu.mx  
**Traductor**

MC. Juan Ulisses Gallardo Zazueta  
**Diseñador y Maquetado**

## Comité Editorial

Dr. Humberto Rodríguez López (UAS)  
humbertorl@uas.edu.mx

Dr. José de Jesús Uriarte Adrian (UAS)  
jesusuriarte@uas.edu.mx

Dr. Juan Francisco Figueroa Pérez (UAS)  
juanfco.figueroa@uas.edu.mx

Dra. Lidia Yadira Pérez Aguilar (UAS)  
lidiaperez@uas.edu.mx

Dra. María del Rosario Salmán Valdez (UAS)  
chayitosalman@uas.edu.mx

Dra. María Guadalupe Soto Decuir (UAS)  
lupita.soto.decuir@uas.edu.mx

Dra. Natividad Cobarrubias Soto (UAS)  
natividadcs@uas.edu.mx

Dr. Ramón Fernando López Osorio (UAS)  
errlop@uas.edu.mx

MC. Raúl Quevedo García (UAS)  
raul.quevedo@info.uas.edu.mx

MC. Thania Roxaana Félix González (UAS)  
thaniafelix@uas.edu.mx

Dr. Topacio Osuna Altamirano (UAS)  
topacio@uas.edu.mx

Dr. Wenseslao Plata Rocha (UAS)  
wenses@uas.edu.mx

## Comité Científico

Dr. Amilcar Meneses Viveros  
Centro de Investigación y de Estudios  
Avanzados del Instituto Politécnico  
Nacional

Dra. Dora Aydee Rodríguez Vega  
Universidad Politécnica de Sinaloa

MC. Elena Muñoz España  
Universidad del Cauca, Colombia

Dra. Graciela Rodríguez Vega  
Universidad de Sonora

Dr. Inés Fernando Vega López  
Universidad Autónoma de Sinaloa

Dr. Juan Antonio Aguilar Rodríguez  
Universidad Autónoma de Occidente

Dr. Juan Cayetano Niebla Zatarain  
Universidad Autónoma de Occidente

Dr. Juan Manuel Ibarra Zannatha  
Centro de Investigación y de Estudios  
Avanzados del Instituto Politécnico  
Nacional

Dr. José Ángel González Fraga  
Universidad Autónoma de Baja California

Dr. Oscar Lozano Carrillo  
Universidad Autónoma Metropolitana

Dr. Pedro Luis Manuel Podesta Lerma  
Universidad Autónoma de Sinaloa

## **Cintillo Legal**

International Journal of Information Science and Technological Applications-UAS IJISTA, es una publicación semestral editada por la Universidad Autónoma de Sinaloa, a través del Cuerpo Académico Realidad Virtual y Robótica (UAS 254) y la Facultad de Informática Culiacán, Dirección: C. Josefa Ortiz de Domínguez S/N, Cd. Universitaria, Ciudad Universitaria, 80013, Culiacán Rosales, Sinaloa, México. Tel. 667 716 1361. <https://revistas.uas.edu.mx/index.php/IJISTA>, editor.ijista@uas.edu.mx. Editora responsable: Dra. Xiomara Penélope Zaldívar Colado. Reservas de Derecho al Uso Exclusivo: 04-2026-033117491800-102, ISSN (en trámite), ambos otorgados por el Instituto Nacional de Derechos de Autor. Las opiniones expresadas por los/las autores/as no reflejan la postura del editor de la publicación. Todos los artículos son de creación original del autor, por lo que esta revista se deslinda de cualquier situación legal derivada por plagios, copias parciales o totales de otros artículos ya publicados y la responsabilidad legal recaerá directamente en el autor del artículo. Cada manuscrito está bajo la licencia Atribución-NoComercial-SinDerivadas 4.0 Internacional (CC BY-NC-ND 4.0) <https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es>

# CONTENIDO

## Editorial

<b>Carta del Editor</b> Dra. Xiomara Penélope Zaldívar Colado.	<b>08 – 09</b>
---	----------------

## Artículos Científicos

<b>Análisis Comparativo del Rendimiento de Inferencia de MobileNetV2 en CPU, GPU y TPU para la Clasificación de Objetos Agrícolas en Sistemas Embebidos</b> J.A. Becerra-Jiménez, J.R. Millán-Almaraz	<b>10 – 18</b>
--	----------------

<b>Propuesta de Aplicación de Monitoreo Remoto para la Automatización de Aulas Orientado al Ahorro Energético en Instituciones Educativas</b> Valeria Jaqueline Gómez-Vázquez, Rossy Nallely Bautista-Vázquez, Ulises Zaldívar Colado	<b>19 – 27</b>
--	----------------

<b>Propuesta de un Visor Geográfico en la ciudad Mazatlán, Sinaloa</b> Michelle Quintero-Bonilla, Yaritza Alejandra Laureano-delaCruz, Luis Jesús Vázquez-Osuna, Zeus Del Valle-Castillo Nájera	<b>28 – 35</b>
--	----------------

<b>Asistente inteligente para consultas académicas y administrativas de la Facultad de Informática de la UAS</b> Cristhian Alexis Ortiz Valentin, Héctor Manuel Padilla Osuna, Héctor Joaquin Escobar Cuevas, Manuel Camacho Martinez	<b>36 – 46</b>
--	----------------

<b>Propuesta de diseño y desarrollo de una aplicación web de autoestudio para estudiantes universitarios: Un enfoque en la gestión del tiempo y hábitos de estudio</b> Arleth de los Ángeles González Ramírez, Valeria Salas Félix, Juan C. Niebla Zatarain	<b>47 – 54</b>
--	----------------

## Editorial

### Carta del Editor

Apreciada comunidad académica, investigadores, lectores y colaboradores de IJISTA:

Es con especial satisfacción que la revista International Journal of Information Science and Technological Applications-UAS IJISTA, fundada por el Cuerpo Académico Realidad Virtual y Robótica y adscrita a la Facultad de Informática Culiacán de la Universidad Autónoma de Sinaloa, presenta el volumen II, número 1, correspondiente a Mayo de 2026. Esta edición reúne contribuciones académicas que documentan el progreso continuo de la investigación aplicada en ciencias de la información, desarrollo tecnológico e innovación computacional.

Los trabajos que conforman este número se articulan alrededor de un propósito común: el diseño, análisis e implementación de soluciones tecnológicas orientadas a resolver necesidades contemporáneas en contextos educativos, institucionales, urbanos y productivos. Desde esta perspectiva, la presente edición subraya la importancia de la informática como disciplina transversal, cuya capacidad para articular fundamentos teóricos con metodologías de desarrollo, validación experimental y aplicación práctica resulta fundamental para el beneficio de la sociedad.

En primer lugar, este número incluye un análisis comparativo del desempeño computacional de la arquitectura MobileNetV2 en configuraciones de procesamiento con CPU, GPU y TPU, aplicada a la clasificación de objetos en contextos agrícolas mediante sistemas embebidos. Esta investigación proporciona evidencia empírica sobre las relaciones entre precisión clasificatoria, velocidad de procesamiento y eficiencia computacional, aspectos fundamentales para el desarrollo de soluciones de agricultura de precisión y tecnologías de visión por computadora aplicables a escenarios operacionales reales.

Asimismo, se incluye una propuesta de sistema de monitoreo remoto para la automatización de espacios educativos, orientada hacia la optimización del consumo energético en instituciones de educación superior. Este estudio examina la viabilidad de tecnologías de Internet de las Cosas (IoT) de bajo costo como alternativa para mejorar la eficiencia en la gestión de recursos energéticos y contribuir al desarrollo de entornos educativos inteligentes y ambientalmente sostenibles.

De manera complementaria, se presenta una propuesta de herramienta cartográfica interactiva para la ciudad de Mazatlán, Sinaloa, cuyo objetivo es facilitar la integración,

---

gestión y visualización de información territorial mediante plataformas geoespaciales de código abierto. Esta contribución adquiere relevancia significativa para el fortalecimiento de procesos de planificación urbana, la toma de decisiones fundamentada en datos espaciales y el acceso equitativo a información de interés público.

En el ámbito de las aplicaciones de inteligencia artificial a la gestión académica, esta edición incorpora el desarrollo de un asistente inteligente para la atención de consultas académicas y administrativas de la Facultad de Informática de la Universidad Autónoma de Sinaloa. Esta propuesta integra técnicas de procesamiento de lenguaje natural, modelos de lenguaje avanzados y estrategias de recuperación aumentada por generación, con el propósito de optimizar la atención institucional y garantizar la disponibilidad de información verificable para la comunidad estudiantil.

Finalmente, se presenta el diseño y conceptualización de EduSelf, una aplicación web dirigida al fortalecimiento del estudio independiente, la organización temporal del aprendizaje y la consolidación de hábitos de estudio efectivos. Este trabajo destaca la relevancia de desarrollar herramientas digitales con enfoque centrado en el usuario, fundamentadas en principios de aprendizaje autorregulado, diseño participativo y apoyo tecnológico para la mejora del desempeño académico.

En conjunto, las contribuciones publicadas en esta edición reflejan la diversidad temática, la solidez metodológica y el compromiso científico de autoras y autores con la generación de conocimiento aplicado de calidad.

Expresamos nuestro más sincero agradecimiento a nuestras autoridades universitarias, así como a todas aquellas personas que participaron en la elaboración, revisión académica y consolidación de esta edición. La contribución de autoras, autores, pares evaluadores y colaboradores institucionales fortalece la calidad editorial de la revista y contribuye al posicionamiento de IJISTA como un espacio de excelencia para la difusión de investigaciones en ciencias de la información y aplicaciones tecnológicas.

Con mis respetos académicos,

**Dra. Xiomara Penélope Zaldívar Colado**  
**Editora en Jefe**  
**International Journal of Information Science and Technological Applications-UAS**  
**Universidad Autónoma de Sinaloa**  
**Mayo de 2026.**

# Análisis Comparativo del Rendimiento de Inferencia de MobileNetV2 en CPU, GPU y TPU para la Clasificación de Objetos Agrícolas en Sistemas Embebidos

## Comparative Analysis of MobileNetV2 Inference Performance on CPU, GPU and TPU for Agricultural Object Classification in Embedded Systems

J.A. Becerra-Jimenez<sup>1</sup>, J.R. Millan-Almaraz<sup>2</sup>

<sup>1</sup>Facultad de Informática / Universidad Autónoma de Sinaloa, México.

<sup>2</sup>Facultad de ciencias Fisicomatemáticas / Universidad Autónoma de Sinaloa, México.

J.A. Becerra-Jimenez, jaimeandres@uas.edu.mx , ORCID: 0009-0007-9816-8010

J.R. Millan-Almaraz, jrmillan@uas.edu.mx , ORCID: 0000-0002-3800-3712

**Recibido:** abril 2026, **Aceptado:** abril 2026, **Publicado:** mayo 2026

### Resumen:

El avance de la agricultura inteligente requiere soluciones de visión por computadora eficientes que operen en dispositivos de bajo costo y consumo energético. Este artículo presenta un análisis comparativo del rendimiento de inferencia de MobileNetV2 en diferentes arquitecturas de hardware: una CPU de escritorio (Ryzen 5 5600G), una GPU (NVIDIA GTX 1070) y un sistema embebido con TPU (Google Coral Dev Board). El estudio evaluó el tiempo de inferencia y los cuadros por segundo (FPS) en modos por lotes y en línea. Los experimentos se realizaron con bases de datos de insectos agrícolas, como la mosquita blanca.

Los resultados muestran que la GPU es la más rápida en inferencia por lotes, alcanzando hasta 683 FPS con 97.93% de precisión. En contraste, la TPU Coral es más eficiente en la inferencia en línea, logrando hasta 348 FPS con 91.76% de precisión. La CPU presenta un rendimiento intermedio y menor eficiencia energética. Estos hallazgos confirman la viabilidad de MobileNetV2 para aplicaciones agrícolas en dispositivos embebidos y demuestran que la elección del hardware depende de la aplicación: la GPU es ideal para el procesamiento masivo, mientras que la TPU es óptima para aplicaciones en tiempo real en el campo.

### Palabras Clave:

*Agricultura de Precisión, MobileNetV2, TPU, GPU, CPU, Sistemas Embebidos.*

### Abstract:

The advancement of smart agriculture requires efficient computer vision solutions that can operate on low-cost, low-power devices. This study presents a comparative analysis of the inference performance of the MobileNetV2 model across different hardware architectures: a desktop CPU (Ryzen 5 5600G), GPU (NVIDIA GTX 1070), and embedded system equipped with a TPU (Google Coral Dev Board). The study evaluated inference time and frames per second (FPS) in both batch and online processing modes using datasets of agricultural insects, such as the whitefly. The results demonstrate that the GPU is the fastest architecture for batch inference, reaching up to 683.28 FPS with a 97.93% accuracy rate. In contrast, the Coral TPU proved to be the most efficient for online inference, achieving up to 348.96 FPS with 91.76% accuracy, whereas the CPU exhibited intermediate performance and lower energy efficiency. These findings confirm the viability of deploying MobileNetV2 for agricultural applications on embedded devices, highlighting that hardware selection depends strictly on the application: GPUs are ideal for massive data processing, while TPUs are optimal for real-time, edge-computing deployments in the field.

### Keywords:

*Precision agriculture, MobileNetV2, TPU, GPU, CPU, Embedded systems.*

## 1. Introducción

La agricultura continúa siendo uno de los pilares fundamentales de la economía de México y de América Latina, y su sostenibilidad depende, en gran medida, de la capacidad del sector para detectar y controlar oportunamente plagas y enfermedades. La mosquita blanca (*Bemisia tabaci* Gennadius) representa uno de los mayores desafíos fitosanitarios a nivel mundial: este insecto chupador, capaz de colonizar más de 600 especies de plantas hospederas, provoca pérdidas económicas de miles de millones de dólares anuales mediante daño directo por succión de savia y, de forma indirecta, actuando como vector de más de 150 virus vegetales, principalmente del género *Begomovirus* [1]. Los impactos documentados son cuantiosos: en el estado de Georgia (EE.UU.), por ejemplo, brotes de *B. tabaci* generaron pérdidas de 132.3 y 161.2 millones de dólares en 2016 y 2017, respectivamente [1]. Su presencia en cultivos hortícolas de la región noroeste de México donde el tomate, el pepino y el chile de agricultura protegida tienen gran relevancia económica convierte el monitoreo temprano y automatizado de esta plaga en una necesidad prioritaria para el sector agroalimentario.

Frente a este escenario, la inteligencia artificial (IA), y en particular el aprendizaje profundo (deep learning), ha emergido como una herramienta de alto potencial para automatizar el reconocimiento de plagas mediante visión por computadora. Su integración en sistemas de Agri-IoT ofrece oportunidades para optimizar la toma de decisiones, mejorar la gestión de recursos fitosanitarios y aumentar la productividad de manera sostenible [2]. Modelos de clasificación de imágenes entrenados con datos agrícolas han demostrado ser capaces de identificar especies de insectos dañinos con alta precisión, incluso bajo condiciones de campo complejas como variaciones de iluminación, fondos no controlados o solapamiento de especímenes [3]. La aplicación de estas técnicas puede transformar el monitoreo tradicional que depende de inspecciones manuales periódicas en sistemas de vigilancia continua y en tiempo real, reduciendo significativamente el tiempo de respuesta ante la aparición de infestaciones.

Sin embargo, el despliegue de estos modelos en entornos agrícolas reales impone restricciones que van más allá de la precisión del modelo. La baja conectividad a Internet en zonas rurales y periurbanas, la limitación en el consumo energético de los dispositivos de campo, y la necesidad de procesamiento en tiempo real hacen que las soluciones basadas en la nube sean frecuentemente inviables o insuficientes [4]. Este contexto ha impulsado el paradigma del cómputo en el borde (edge computing), en el cual el procesamiento ocurre directamente en el dispositivo ubicado en el campo, sin necesidad de transmitir datos a servidores remotos. Esta arquitectura ofrece ventajas claras en cuanto a latencia, privacidad de los datos y autonomía operativa [5]. No obstante, su

implementación plantea una pregunta de diseño crítica: ¿cuál plataforma de hardware embebido ofrece el mejor equilibrio entre velocidad de inferencia, precisión del modelo y viabilidad de despliegue para aplicaciones agrícolas específicas? Responder esta pregunta requiere una evaluación experimental sistemática, dado que los dispositivos disponibles desde CPU de propósito general hasta aceleradores de IA especializados presentan perfiles de rendimiento y consumo muy distintos entre sí [6].

En este contexto, los modelos de clasificación ligeros han demostrado ser los candidatos naturales para el despliegue en sistemas embebidos. Entre ellos, MobileNetV2, propuesto por Sandler et al. [7], destaca por su arquitectura basada en residuales invertidos y cuellos de botella lineales (inverted residuals and linear bottlenecks), que permite operar con aproximadamente 300 millones de operaciones de punto flotante (FLOPs) y tan solo 3.4 millones de parámetros, manteniendo una precisión competitiva en tareas de clasificación de imágenes. A diferencia de otras alternativas ligeras, como MobileNetV3 o EfficientNet que, si bien son más precisas en ciertos benchmarks, requieren la función de activación hard-swish que resulta incompatible con el compilador de la Edge TPU de Google Coral [6], MobileNetV2 ofrece compatibilidad nativa y completa con dicha plataforma de hardware, permitiendo que todas sus operaciones sean ejecutadas íntegramente en el acelerador sin necesidad de recurrir a la CPU del sistema embebido. Estudios comparativos recientes han confirmado que, en escenarios de recursos limitados, MobileNetV2 demuestra una superior compatibilidad con hardware de borde y velocidades de inferencia más consistentes frente a EfficientNetV2 [8]. Su eficacia en aplicaciones agrícolas ha sido validada en múltiples trabajos, incluyendo la detección de enfermedades en hojas de tomate y la clasificación de plantas en condiciones de campo [3].

El presente estudio se enfoca en la evaluación y comparación del rendimiento de inferencia de MobileNetV2 en tres arquitecturas de hardware representativas del cómputo tradicional y del edge computing: una CPU de escritorio (AMD Ryzen 5 5600G), una GPU de consumo (NVIDIA GTX 1070) y un sistema embebido con coprocesador TPU (Google Coral Dev Board). El objetivo central es determinar cuál de estas plataformas resulta más adecuada para clasificar insectos agrícolas con énfasis en la mosquita blanca (*B. tabaci*) bajo dos modalidades operativas: inferencia en línea, que simula el monitoreo en tiempo real imagen por imagen, e inferencia por lotes, orientada al análisis masivo de colecciones de imágenes. Las métricas evaluadas son el tiempo promedio de inferencia, los cuadros por segundo (FPS) y la precisión del modelo. Los resultados buscan proveer criterios objetivos que orienten la selección de hardware para sistemas de visión embebida en agricultura de precisión, especialmente en contextos con restricciones de infraestructura tecnológica y energética, y sentar bases

experimentales para el desarrollo de soluciones de Agri-IoT de bajo costo desplegadas en zonas rurales.

## 2. Trabajos Relacionados

La investigación presentada en este artículo se sitúa en la intersección de tres áreas de conocimiento activas, el primero es la aplicación de inteligencia artificial e IoT en la agricultura de precisión, el segundo es uso de modelos ligeros de visión por computadora para la detección de plagas y enfermedades en cultivos y, el tercero la comparación experimental de plataformas de hardware para inferencia en el borde. A continuación, se revisan los trabajos más relevantes en cada uno de estos ejes temáticos, identificando los vacíos que el presente estudio busca llenar.

### 2.1. Inteligencia artificial e IoT en la agricultura de precisión

La integración de tecnologías como el Internet de las Cosas (IoT), la inteligencia artificial (IA) y los sistemas embebidos ha transformado profundamente la gestión de las operaciones agrícolas en lo que se conoce como Agri-IoT. Esta convergencia tiene un potencial significativo para optimizar la toma de decisiones, mejorar la gestión de recursos hídricos y fitosanitarios, y aumentar la productividad de forma sostenible. Lykas y Vagelas resaltan que la adopción de estas tecnologías en sistemas agroalimentarios no solo mejora la eficiencia productiva, sino que también contribuye a la sostenibilidad ambiental a largo plazo, subrayando la relevancia de innovar en sistemas embebidos para el monitoreo en tiempo real [2].

En cuanto al monitoreo ambiental y de recursos, Morchid et al. propusieron un sistema de riego inteligente basado en IoT para la gestión del agua agrícola que integra sistemas embebidos, telemetría y cómputo en la nube, demostrando la viabilidad de soluciones de bajo costo para decisiones de campo en tiempo real [5]. Por su parte, Dos Santos et al. desarrollaron un sistema embebido que combina redes de sensores inalámbricos con un motor de predicción basado en el modelo ARIMA para anticipar condiciones adversas en cultivos, mostrando que la arquitectura de procesamiento local reduce la latencia de respuesta frente a enfoques que dependen exclusivamente de la nube [9].

La aplicación de algoritmos de aprendizaje automático (ML) y técnicas de IA en la agricultura ha sido ampliamente documentada como vía para optimizar la producción, predecir tendencias, identificar patrones y automatizar tareas críticas [10]. Saez Rojas et al. presentaron un prototipo IoT para la optimización del riego agrícola en la Región del Biobío (Chile), donde la combinación de sensores en campo con modelos predictivos permitió reducir el consumo de agua sin afectar el rendimiento de los cultivos [11]. El

denominador común en todos estos estudios es la necesidad de soluciones que sean no solo precisas, sino también desplegadas en condiciones de infraestructura limitada, lo que motiva directamente el enfoque de edge computing adoptado en el presente trabajo.

Pintus et al. documentaron la tendencia hacia sistemas de Edge AIoT en tiempo real para la clasificación de imágenes agrícolas, argumentando que los avances en hardware acelerador de IA y en modelos de aprendizaje profundo han hecho factible el procesamiento directo en el dispositivo de campo, eliminando la dependencia de la nube para tareas de detección urgente [6][4]. Esta transición es particularmente relevante para regiones con conectividad a Internet intermitente o inexistente, como las zonas agrícolas rurales del noroeste de México.

### 2.2 Modelos ligeros de visión por computadora para la detección de plagas y enfermedades

El desarrollo de modelos de visión por computadora aplicados a la detección de plagas y enfermedades agrícolas ha avanzado de forma significativa en los últimos años, impulsado en gran medida por la disponibilidad de arquitecturas CNN ligeras diseñadas para dispositivos con restricciones de cómputo. En este contexto, MobileNetV2 ha emergido como una de las arquitecturas de referencia más utilizadas, gracias a su eficiente balance entre precisión y requisitos computacionales [7].

Sharma et al. demostraron la eficacia de MobileNetV2 en la clasificación de enfermedades en hojas de tomate mediante un modelo ensemble que combina esta arquitectura con ResNet50, alcanzando una exactitud del 99.91% sobre un conjunto de datos de 11,000 imágenes en 10 categorías [3]. Este resultado refuerza la capacidad de MobileNetV2 para extraer características relevantes en imágenes agrícolas complejas incluso con tamaños de conjunto de datos moderados, lo cual es particularmente valioso en escenarios de monitoreo regional donde la recopilación masiva de imágenes resulta costosa.

Para el caso específico de detección de insectos, la revisión sistemática de Teixeira et al. sobre la detección automática de insectos mediante aprendizaje profundo documenta múltiples estudios que emplean trampas amarillas adhesivas e imágenes de campo para la identificación y conteo de plagas como mosquita blanca, áfidos y trips [12]. Los autores señalan que la variabilidad visual de los insectos en condiciones de campo determinada por factores como la distancia de captura, la densidad de organismos y las condiciones de iluminación constituye el principal reto técnico para la generalización de estos modelos, lo que justifica el uso de conjuntos de datos diversificados como el adoptado en el presente trabajo.

Srinivasa [13] y otros trabajos orientados al uso de drones para la detección temprana de plagas y malezas en

campo destacan la ventaja del monitoreo aéreo automatizado, pero señalan que el procesamiento en tiempo real de las imágenes sigue siendo un cuello de botella cuando se depende de transmisión a la nube, reforzando la necesidad de soluciones de inferencia embebida. Dong et al. propusieron un sistema IoT de monitoreo ambiental agrícola que combina LoRaWAN para transmisión de largo alcance con un módulo de reconocimiento de plagas basado en TensorFlow ejecutado en nodos de borde; reportaron una precisión de reconocimiento del 89% con tiempos de procesamiento compatibles con la operación en campo, y concluyeron que el desplazamiento del cómputo hacia el borde reduce en más del 60% la carga de transmisión y procesamiento en la nube [14].

La literatura revisada muestra consistentemente que MobileNetV2 es una de las arquitecturas con mayor adopción en sistemas agrícolas embebidos, en parte porque fue diseñada explícitamente para operar en dispositivos móviles y de baja potencia [7], y en parte porque su compatibilidad nativa con TensorFlow Lite facilita su conversión y despliegue en plataformas de hardware especializado como la Edge TPU de Google Coral [6]. Sin embargo, hasta donde se ha identificado, pocos estudios han realizado una comparación experimental directa y sistemática del rendimiento de inferencia de un mismo modelo en CPU de escritorio, GPU y TPU bajo modalidades de procesamiento en línea y por lotes, en el dominio específico de insectos agrícolas, lo que representa el aporte central del presente artículo.

### 2.3 Comparación de hardware para inferencia en el borde

La selección del hardware adecuado para la inferencia de modelos de aprendizaje profundo en sistemas embebidos es una decisión de diseño crítica con implicaciones directas sobre la latencia, el consumo energético y el costo de implementación. Diversos estudios han abordado esta pregunta comparando las plataformas disponibles en el mercado de edge computing, con resultados que orientan el diseño de los experimentos del presente artículo.

Tobiasz et al. realizaron un estudio comparativo de inferencia en plataformas de edge computing, Google Coral USB, Google Coral PCIe, Intel Neural Compute Stick 2 y NVIDIA Jetson Nano. Evaluando las familias MobileNet y EfficientNet con diferentes tamaños de entrada [6]. Sus resultados confirman que, para MobileNetV2 con resolución de entrada  $224 \times 224$ , los dispositivos Coral logran el mayor rendimiento en FPS, superando a Jetson Nano en un factor de  $4.42 \times$  y al Neural Compute Stick 2 en  $9.08 \times$ . Adicionalmente, documentan que MobileNetV3 no puede ejecutarse íntegramente en la Edge TPU por incompatibilidad de la función de activación hard-swish, lo que explica la selección de MobileNetV2 en el presente estudio.

Pérez-García et al. ampliaron este análisis comparando cinco dispositivos de edge computing los cuales eran la Raspberry Pi 4, Google Coral Dev Board, Google Coral Mini, NVIDIA Jetson Nano y HummingBoard Pro. Utilizando múltiples modelos de clasificación y detección de objetos [15]. Midieron tiempos de inferencia, consumo de RAM, uso de CPU y energía, encontrando que la Edge TPU de Coral logra la mejor eficiencia energética durante la inferencia activa, mientras que el Jetson Nano consume significativamente más RAM. Los autores concluyen que la selección del dispositivo óptimo depende del perfil de uso: Coral es superior para inferencia continua de alta frecuencia, mientras que Jetson ofrece mayor flexibilidad para modelos que no pueden cuantizarse completamente a INT8.

En el ámbito específico de la agricultura, la revisión de Pintus et al. sobre edge AIoT para clasificación de imágenes agrícolas en tiempo real analiza el estado del arte en hardware acelerador de IA incluyendo GPU, TPU, NPU y FPGA y concluye que el flujo de trabajo óptimo para sistemas Agri-IoT consiste en entrenar y optimizar el modelo de aprendizaje profundo en hardware de alto rendimiento (GPU), y, convertir y desplegar el modelo en dispositivos de edge para la inferencia en campo [6]. Este flujo de dos etapas es precisamente el adoptado en el presente estudio. Adicionalmente, los autores identifican como uno de los principales desafíos vigentes la falta de benchmarks estandarizados que permitan comparar plataformas de hardware bajo condiciones de aplicación agrícola específicas, señalando así el vacío que esta investigación contribuye a llenar.

En conjunto, los trabajos revisados en esta sección muestran que, si bien existen comparaciones de hardware para inferencia de modelos ligeros, estas se realizan generalmente con datasets de propósito general (ImageNet, CIFAR) y no con colecciones de imágenes de plagas agrícolas capturadas en condiciones de campo real. El presente estudio se distingue por evaluar el rendimiento de inferencia sobre un dataset propio de insectos agrícolas de importancia fitosanitaria para el noroeste de México, añadiendo la perspectiva de la modalidad de procesamiento (en línea vs. por lotes) como variable de análisis, lo que aporta directrices prácticas para el diseño de sistemas de monitoreo en este dominio.

### 3. Metodología

Para el análisis, se utilizó el modelo MobileNetV2, conocido por su eficiencia y diseño optimizado para dispositivos móviles y embebidos. El entrenamiento del modelo se realizó con un conjunto de datos que contenía 250 imágenes para entrenamiento y 51 para validación por cada clase, incluyendo la mosca blanca, mosquitos, trips y otros insectos. La Tabla 1 resume la distribución del conjunto de datos utilizado para el entrenamiento del modelo.

Tabla 1. Fotos usadas en entrenamiento

Clase	Fotos entrenamiento	Fotos validación
Mosquita blanca	250	51
Mosquitos	250	51
Trip	250	51
Otros insectos	250	51

El resultado del entrenamiento de este conjunto de imágenes confirmó la convergencia del modelo y su alta precisión, obteniendo las siguientes métricas:

- Precisión: 0.9967
- Pérdida: 0.006
- Precisión de Validación: 0.9892
- Pérdida de Validación: 0.0892

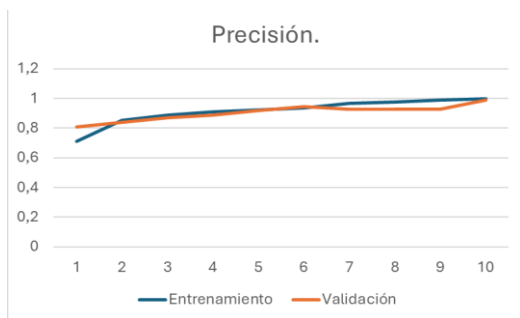


Fig. 1. Gráfica de precisión en el entrenamiento

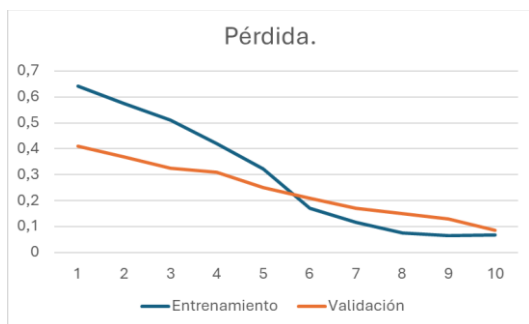


Fig. 2. Gráfica de pérdida en el entrenamiento

Para diversificar los casos y asegurar la robustez del modelo al enfrentarse a diferentes escenarios en el campo, se utilizó un conjunto de datos variado para la inferencia. Las imágenes capturan ejemplos de la mosquita blanca en distintos contextos, tales como sobre una hoja de planta

(ver Fig. 3), en trampas amarillas a mayor distancia (ver Fig. 4) y acompañada de otros insectos en tomas de mayor acercamiento (ver Fig. 5).



Fig.3. Mosquita blanca en hoja

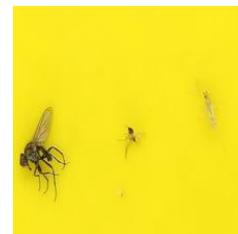


Fig. 4. Mosquita blanca en trampa amarilla



Fig. 5. Mosquita blanca en trampa amarilla

Una vez entrenado el modelo, el experimento se llevó a cabo en las siguientes plataformas de hardware para realizar las pruebas de inferencia:

PC de Escritorio: Equipada con una CPU AMD Ryzen 5 5600G y una GPU NVIDIA 1070.

Sistema Embebido: Una Coral Dev Board con su CPU NXP i.MX 8M SoC y el coprocesador Edge TPU ML accelerator.

Se realizaron pruebas de inferencia utilizando un conjunto de datos propio (ver Tabla 2).

Tabla 2. Fotos para inferencias.

Clase	Fotos para inferencias
Mosquita blanca	75
Mosquitos	52
Trip	38
Otros insectos	55

Las pruebas se ejecutaron bajo dos modalidades:

- Modo en línea (*Online*): Se procesó una sola imagen por cada iteración,

simulando una aplicación en tiempo real.

- Modo por lotes (Batch): Se procesaron múltiples imágenes en un solo lote, lo cual es relevante para tareas que requieren un alto rendimiento masivo, como el análisis de imágenes aéreas.

Las métricas evaluadas fueron el tiempo promedio de inferencia, los cuadros por segundo (FPS) y el porcentaje de aciertos del modelo. Es importante destacar que no fue posible realizar la inferencia por lotes en el núcleo TPU debido a su arquitectura, ya que su entrada solo permite inferir una imagen a la vez.

#### 4. Resultados

A continuación, se presenta la comparativa de los resultados de inferencia obtenidos en las distintas plataformas y modalidades evaluadas. La Tabla 3 resume las métricas de tiempo promedio de inferencia, cuadros por segundo (FPS) y porcentaje de precisión del modelo MobileNetV2. Es importante destacar que no fue posible realizar la inferencia por lotes en el núcleo TPU debido a su arquitectura, ya que su entrada solo permite inferir una imagen a la vez.

Tabla 3. comparativa de inferencias en diferentes escenarios

Tipo de inferencia	Dispositivo	Tiempo prom. inferencia (s)	FPS	Precisión (%)
Lotes	ARM-Cortex A53	0.2129	4.69	96.50%
	Ryzen 5600	0.01286	72.13	97.93%
	NVIDIA 1070	0.00146	683.28	97.93%
Online	Ryzen 5600	0.0464	21.54	96.55%
	NVIDIA 1070	0.0415	23.93	96.5%
	ARM-Cortex A53	0.4217	2.37	91,13%

	TPU Coral	0.00287	348.96	91.76%
--	-----------	---------	--------	--------

Para una visualización más clara de la tasa de procesamiento, se generaron gráficas de barras con desviación estándar para cada modalidad. Como se observa en la Fig. 6, en el modo por lotes, la GPU NVIDIA 1070 presenta la mayor cantidad de cuadros por segundo.

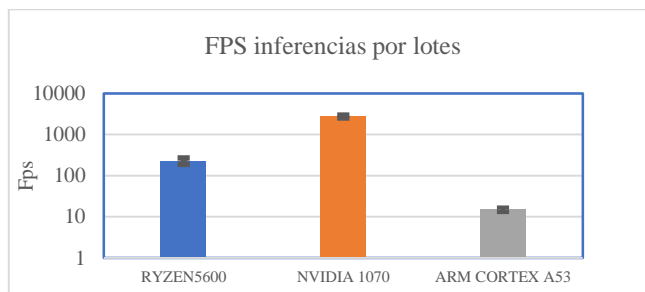


Fig. 6. Comparativa FPS inferencias por lotes

Por otro lado, la Fig. 7 muestra el rendimiento en FPS en la modalidad en línea, destacando las métricas obtenidas por la TPU Coral frente al resto de los dispositivos.

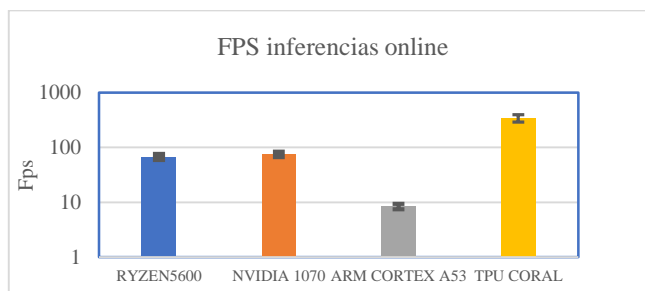


Fig. 7. Comparativa FPS inferencias online

Para registrar el comportamiento del hardware a través del tiempo, se graficó el tiempo total requerido para cada iteración. En la modalidad de inferencia por lotes, la Fig. 8 muestra el desempeño en la CPU del sistema embebido (ARM-Cortex A53). Por su parte, la Fig. 9 presenta los tiempos de la CPU Ryzen 5600, y la Fig. 10 ilustra la inferencia en la GPU NVIDIA 1070.

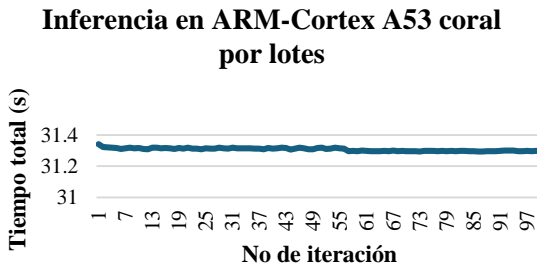


Fig. 8. Inferencia ARM-Cortex A53 coral por lotes

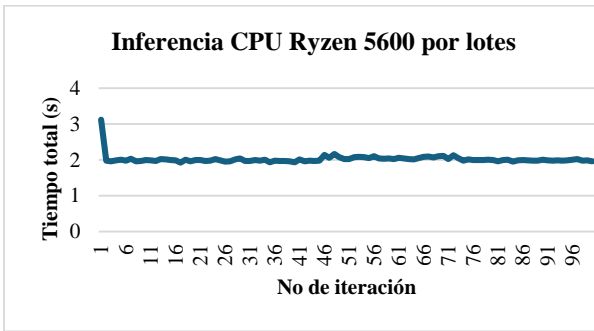


Fig. 9. Inferencia CPU Ryzen 5600 por lotes

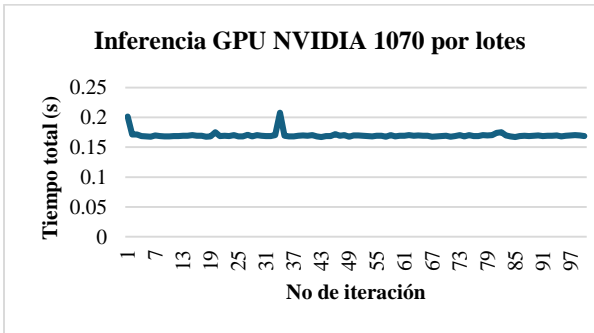


Fig. 10. Inferencia GPU NVIDIA 1070 por lotes

En cuanto a la modalidad en línea, la Fig. 11 detalla la curva de iteraciones de la CPU Ryzen 5600. La Fig. 12 expone el rendimiento de la GPU NVIDIA 1070. La Fig. 13 evidencia los tiempos de inferencia de la CPU del sistema embebido y, finalmente, la Fig. 14 documenta el comportamiento de la TPU Coral a lo largo de las cien iteraciones.

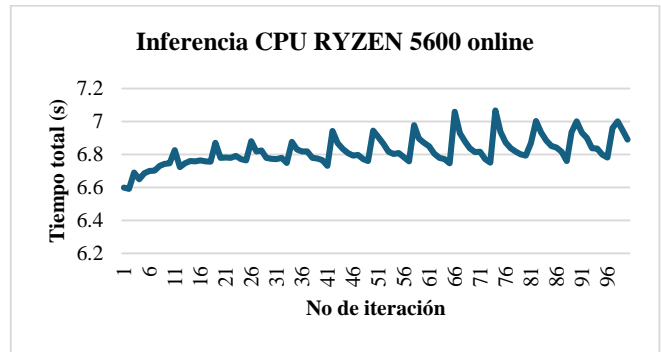


Fig. 11 Inferencia CPU Ryzen 5600 online

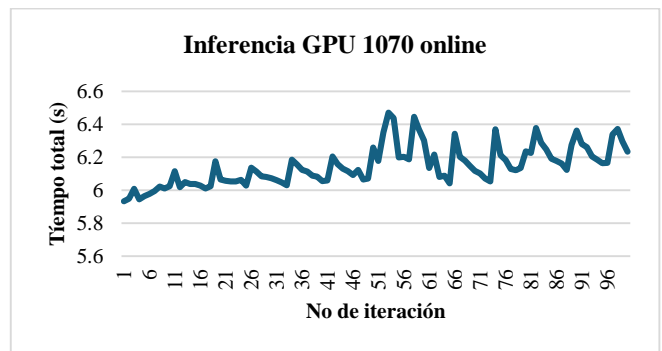


Fig. 12. Inferencia GPU 1070 online

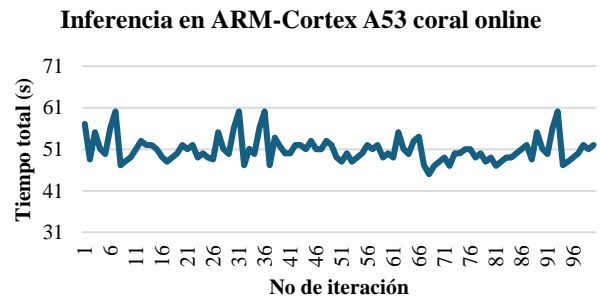


Fig. 13. Inferencia en ARM-Cortex A53 coral online

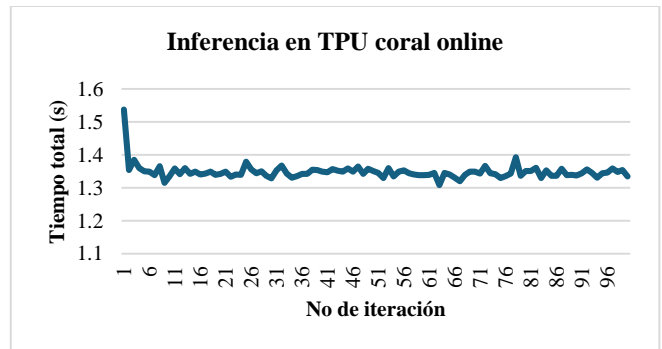


Fig. 14. Inferencia en TPU coral online

## 5. Análisis de Resultados

Los datos obtenidos confirman que el rendimiento de inferencia depende de manera directa de la arquitectura del hardware utilizado y de la modalidad de procesamiento exigida por la aplicación. La precisión del modelo se mantuvo consistentemente alta en todos los dispositivos, validando la robustez de MobileNetV2; sin embargo, los tiempos de respuesta variaron drásticamente.

En el procesamiento por lotes, la GPU (NVIDIA 1070) demostró una superioridad absoluta al alcanzar el mejor rendimiento con un tiempo de inferencia de 0.00146 segundos y 683.28 FPS. Este resultado se explica por la capacidad inherente de la GPU para paralelizar el procesamiento. Como se evidenció en la Fig. 10, la estabilidad de los tiempos a lo largo de las iteraciones confirma que esta arquitectura es la idónea para tareas que implican el análisis masivo de datos o bases de imágenes recopiladas previamente.

En contraste, para la inferencia en línea la cual es esencial para aplicaciones en tiempo real donde cada imagen debe ser procesada individualmente con la mínima latencia posible, la TPU Coral demostró ser la arquitectura más eficiente. Logró el tiempo de inferencia más bajo en esta modalidad (0.00287 segundos, equivalente a 348.96 FPS). Aunque la GPU también reportó tiempos bajos en la modalidad en línea (Fig. 12), mostró ligeras variaciones y no alcanzó el nivel de optimización de la TPU. La curva casi plana observada en la Fig. 14 confirma que la TPU está diseñada específicamente para ejecutar inferencias con una eficiencia extrema y mínima latencia, cumpliendo con los requisitos críticos del edge computing.

Respecto al rendimiento de las unidades centrales de procesamiento, la CPU de escritorio (Ryzen 5600) mantuvo un desempeño intermedio y estable, aunque con una latencia considerable para aplicaciones de respuesta instantánea. Por último, la CPU del sistema embebido (ARM-Cortex A53) presentó los tiempos de inferencia más altos (0.4217 segundos y 2.37 FPS en modo en línea). Este bajo desempeño subraya que las CPU tradicionales de bajo consumo tienen fuertes limitaciones para el procesamiento de modelos de aprendizaje profundo en tiempo real, lo que hace indispensable la integración de coprocesadores especializados, como la TPU, para viabilizar el desarrollo de sistemas agrícolas inteligentes en dispositivos embebidos.

## 6. Conclusiones

Este estudio concluye que, para la ejecución eficiente de modelos de aprendizaje profundo ligero como MobileNetV2 en aplicaciones de agricultura inteligente, la selección de la arquitectura de hardware es un factor determinante. Los resultados demuestran que la TPU

(Google Coral Dev Board) es la solución óptima para la inferencia en tiempo real en sistemas embebidos, gracias a su excelente rendimiento de latencia y bajo consumo de energía, características indispensables para el edge computing en zonas de cultivo. Por otro lado, para tareas que no exigen inmediatez, como el análisis masivo fuera de línea de grandes bases de datos fotográficas, las GPUs (como la NVIDIA 1070) continúan siendo la opción más eficiente debido a su sobresaliente capacidad de procesamiento paralelo.

La importancia de este análisis radica en demostrar la viabilidad de utilizar dispositivos de bajo costo y bajo consumo energético para tareas críticas, como la detección oportuna de la mosquita blanca, superando las limitaciones tradicionales de las CPU en sistemas embebidos. Como trabajo futuro, se propone desarrollar un prototipo de sistema completo que integre la Coral Dev Board con una cámara para la clasificación automática de objetos agrícolas en un entorno real. Este prototipo permitirá la validación de la metodología directamente en el campo y facilitará su integración con otros componentes del Internet de las Cosas (IoT), sentando las bases para una gestión proactiva, autónoma y más eficiente en el ámbito de la agricultura de precisión.

## 7. Referencias

- [1] Y. Li, G. N. Mbata, S. Punnuri, A. M. Simmons, y D. I. Shapiro-Ilan, "Bemisia tabaci on Vegetables in the southern United States: Incidence, impact, and management", *Insects*, vol. 12, núm. 3, pp. 1–29, mar. 2021, doi: 10.3390/insects12030198.
- [2] C. Lykas y I. Vagelas, "Innovations in Agriculture for Sustainable Agro-Systems", el 1 de septiembre de 2023, Multidisciplinary Digital Publishing Institute (MDPI). doi: 10.3390/agronomy13092309.
- [3] J. Sharma *et al.*, "Deep learning based ensemble model for accurate tomato leaf disease classification by leveraging ResNet50 and MobileNetV2 architectures", *Sci. Rep.*, vol. 15, núm. 1, p. 13904, dic. 2025, doi: 10.1038/s41598-025-98015-x.
- [4] M. Pintus, F. Colucci, y F. Maggio, "Emerging Developments in Real-Time Edge AIoT for Agricultural Image Classification", el 1 de marzo de 2025, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/iot6010013.
- [5] A. Morchid, R. Jebabra, H. M. Khalid, R. El Alami, H. Qjidaa, y M. Ouazzani Jamil, "IoT-based smart irrigation management system to enhance agricultural water security using embedded systems, telemetry data, and cloud computing", *Results in Engineering*, vol. 23, p. 102829, sep. 2024, doi: 10.1016/j.rineng.2024.102829.

- [6] R. Tobiasz, G. Wilczynski, P. Graszka, N. Czechowski, y S. Luczak, “Edge Devices Inference Performance Comparison”, *Journal of Computing Science and Engineering*, vol. 17, núm. 2, pp. 51–59, 2023, doi: 10.5626/JCSE.2023.17.2.51.
- [7] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, y L.-C. Chen, “MobileNetV2: Inverted Residuals and Linear Bottlenecks”, en 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 4510–4520. doi: 10.1109/CVPR.2018.00474.
- [8] T. Njoroge, R. Kibuku, y K. Mugoye, “Comparative and edge-hybrid modeling of EfficientNetV2 and MobileNetV2 for multi-classcrop disease classification with statistical validation”, *Journal of Edge Computing*, vol. 4, pp. 234–262, nov. 2025, doi: 10.55056/jec.905.
- [9] U. J. L. dos Santos, G. Pessin, C. A. da Costa, y R. da Rosa Righi, “AgriPrediction: A proactive internet of things model to anticipate problems and improve production in agricultural crops”, *Comput. Electron. Agric.*, vol. 161, pp. 202–213, jun. 2019, doi: 10.1016/j.compag.2018.10.010.
- [10] S. Prof, B. Sharon, P. P, P. S, S. Prince, y R. R, “Smart Agriculture with IoT”, *International Journal of Innovative Research in Information Security*, vol. 09, núm. 03, pp. 225–228, jun. 2023, doi: 10.26562/ijiris.2023.v0903.31.
- [11] J. I. Saez Rojas, J. M. Pantoja, M. Matamala, I. C. Briceno, J. P. Vasconez, y A. R. Romero-Conrado, “An IoT-Based Prototype for Optimizing Agricultural Irrigation: A Case Study in the Biobio Region of Chile”, *Procedia Comput. Sci.*, vol. 238, pp. 1009–1014, 2024, doi: 10.1016/j.procs.2024.06.127.
- [12] A. C. Teixeira, J. Ribeiro, R. Morais, J. J. Sousa, y A. Cunha, “A Systematic Review on Automatic Insect Detection Using Deep Learning”, el 1 de marzo de 2023, *MDPI*. doi: 10.3390/agriculture13030713.
- [13] B. Revanasiddappa, C. S. Arvind, y S. Swamy, “Real-time early detection of weed plants in pulse crop field using drone with IoT”, *International Journal of Agricultural Technology*, vol. 16, núm. 5, pp. 1227–1242, 2020, Consultado: el 4 de mayo de 2026. [En línea]. Disponible en: <https://li04.tci-thaijo.org/index.php/IJAT/article/view/7422>
- [14] M. Dong, H. Yu, Z. Sun, L. Zhang, Y. Sui, y R. Zhao, “Research on agricultural environmental monitoring Internet of Things based on edge computing and deep learning”, *Journal of Intelligent Systems*, vol. 33, may 2024, doi: 10.1515/jisys-2023-0114.
- [15] A. Garcia-Perez, R. Miñón, A. I. Torre-Bastida, y E. Zulueta-Guerrero, “Analysing Edge Computing Devices for the Deployment of Embedded AI”, *Sensors*, vol. 23, núm. 23, p. 9495, dic. 2023, doi: 10.3390/s23239495.

# Propuesta de Aplicación de Monitoreo Remoto para la Automatización de Aulas Orientado al Ahorro Energético en Instituciones Educativas

## Proposal of Remote Monitoring Application Classroom Automation Oriented Toward Energy Savings in Educational Institutional

Valeria Jaqueline Gómez-Vazquez<sup>1</sup>, Rossy Nallely Bautista-Vazquez<sup>1</sup>, Ulises Zaldívar Colado<sup>1</sup>

<sup>1</sup>Facultad de Informática Mazatlán, Universidad Autónoma de Sinaloa, México.

Valeria Jaqueline Gomez-Vazquez, 20857373.gomez@ms.uas.edu.mx, ORCID: 0009-0006-4011-4610

Rosy Nalley Bautista Vazquez, rossybautista167@gmail.com, ORCID: 0009-0003-3143-2279

Autor por correspondencia: Ulises Zaldívar Colado, uzaldivar@uas.edu.mx ORCID 0000-0001-6493-3553

**Recibido:** abril 2026, **Aceptado:** abril 2026, **Publicado:** mayo 2026

### Resumen:

Este trabajo presenta el diseño y desarrollo de una propuesta de sistema de automatización de aulas orientado al ahorro energético mediante el uso de tecnologías IoT de bajo costo. La propuesta integra un microcontrolador ESP32, sensores de presencia y detección de gas, así como actuadores controlados tanto de manera autónoma como remota a través de una aplicación web. La metodología incluyó el diseño de la arquitectura del sistema, su simulación en Tinkercad, la implementación en una maqueta a escala funcional y pruebas de validación en distintos escenarios de operación.

Los resultados obtenidos en el entorno de prueba demuestran que la propuesta es capaz de responder en tiempo real ante cambios en el entorno, activando o desactivando dispositivos eléctricos de manera eficiente. El análisis realizado sugiere la viabilidad de la propuesta como una alternativa accesible y escalable para instituciones educativas con recursos limitados, contribuyendo a la optimización del consumo energético y a la sostenibilidad ambiental.

### Palabras Clave:

*Monitoreo remoto, IoT, eficiencia energética, aulas inteligentes, ESP32, automatización*

### Abstract:

This paper presents the design and development of a proposed classroom automation system aimed at energy savings through the use of low-cost IoT technologies. The proposal integrates an ESP32 microcontroller, presence and gas detection sensors, as well as actuators controlled both autonomously and remotely through a web application. The methodology included the design of the system architecture, its simulation in Tinkercad, implementation in a functional-scale prototype, and validation tests under different operating scenarios.

The results obtained in the testing environment demonstrate that the proposed system is capable of responding in real time to environmental changes by efficiently activating or deactivating electrical devices. The analysis suggests the feasibility of the proposal as an accessible and scalable alternative for educational institutions with limited resources, contributing to energy consumption optimization and environmental sustainability.

### Keywords:

*Remote monitoring, IoT, energy efficiency, smart classrooms, ESP32, automation*

## 1. Introducción

La demanda energética en instituciones educativas ha evidenciado un uso inadecuado de recursos como iluminación, ventilación y equipos eléctricos, generando costos elevados y un impacto ambiental significativo. En numerosas instituciones, las aulas permanecen con luces encendidas aun cuando están desocupadas, aires acondicionados operando sin necesidad o puertas abiertas que afectan la climatización interna. Esta problemática no solo incrementa el gasto económico institucional, sino que también contribuye a una mayor huella de carbono. Ante este escenario, surge la necesidad de implementar soluciones que permitan optimizar el uso de energía en los espacios educativos mediante sistemas inteligentes autónomos.

La administración en las aulas depende de una supervisión humana presencial y reactiva, lo cual limita la capacidad de respuesta ante descuidos, lo que incrementa la probabilidad de descuidos operativos, dando como resultado un desperdicio energético constante y deterioro prematuro de equipos eléctricos. Por ello, resulta necesario diseñar propuestas que integren automatización para regular el consumo de energía de manera eficiente y sostenible. La incorporación de tecnologías inteligentes permitiría no solo reducir costos, sino también fomentar una cultura institucional orientada a la eficiencia energética y la responsabilidad ambiental.

La evolución de los entornos inteligentes puede clasificarse mediante taxonomías de automatización que incluyen sistemas reactivos, sistemas programados y sistemas autónomos adaptativos. Los sistemas reactivos responden a estímulos inmediatos, como sensores de movimiento para iluminación; los programados operan bajo horarios predefinidos; mientras que los adaptativos emplean datos ambientales para ajustar su funcionamiento dinámicamente. En el ámbito educativo, estas categorías han sido aplicadas progresivamente dentro del concepto de edificios inteligentes.

A nivel internacional, el concepto de Smart Campus se ha consolidado como un modelo basado en la integración de tecnologías digitales, sensores e Internet de las Cosas (IoT) para la gestión eficiente de los recursos universitarios. Este enfoque se define como un ecosistema inteligente y sostenible que permite optimizar procesos mediante el uso de plataformas tecnológicas y sistemas automatizados. En este sentido, los Smart Campus no constituyen una implementación específica de una sola institución, sino un marco adaptable que puede ser adoptado por distintas universidades para mejorar la eficiencia energética, la gestión de infraestructuras y la calidad de los servicios. Además, la Universidad Nacional de Educación a Distancia (UNED) en España ha incorporado sensores para monitorear variables como ocupación, temperatura y luminosidad, permitiendo la regulación automática de sistemas de climatización e iluminación en tiempo real, lo que contribuye a la reducción del consumo energético. Asimismo, estos sistemas integran plataformas de gestión centralizada que combinan sensores, automatización y análisis de datos para mejorar la eficiencia operativa de los campus universitarios. Estas iniciativas se enmarcan dentro de los sistemas autónomos adaptativos, ya que ajustan su funcionamiento en función de las condiciones ambientales y

el uso de los espacios, promoviendo un enfoque integral de eficiencia energética institucional [1, 2].

Sin embargo, muchos de los proyectos existentes presentan limitaciones asociadas a altos costos de implementación y dependencia de infraestructura compleja, lo que dificulta su replicabilidad en instituciones con recursos limitados. Algunos modelos priorizan la innovación tecnológica sin considerar escalabilidad o accesibilidad económica.

Bajo este contexto, el presente trabajo describe el desarrollo de un sistema de automatización basado en el microcontrolador ESP32, orientado a la eficiencia energética en aulas mediante un esquema híbrido de sensores autónomos y control remoto. La propuesta busca mantener un bajo costo de implementación mediante el uso de componentes accesibles, estimándose su viabilidad tanto a nivel de un aula individual como en escenarios escalables, como instituciones con múltiples aulas, aspecto que será analizado en secciones posteriores.

## 2. Trabajos Relacionados

El análisis de trabajos relacionados permite contextualizar el desarrollo de sistemas de gestión energética dentro de entornos educativos inteligentes, destacando las principales tecnologías, enfoques y resultados obtenidos en investigaciones previas. En particular, la integración de Internet de las Cosas (IoT), sensores inteligentes y arquitecturas de bajo costo ha impulsado significativamente la optimización del consumo energético en aulas y edificios. A través de la revisión de la literatura, se identifican tendencias clave como la automatización basada en ocupación, el monitoreo en tiempo real y el uso de sistemas contextuales para la toma de decisiones. En las siguientes subsecciones se presentan los principales avances en estas áreas, organizados según su enfoque tecnológico y su aplicación en la eficiencia energética.

### 2.1. Sistemas de gestión energética en aulas inteligentes

Los sistemas de gestión energética (EMS, Energy Management System) han sido ampliamente estudiados dentro del contexto de aulas y campus inteligentes. Estos sistemas buscan optimizar el consumo energético mediante el uso de tecnologías como IoT y redes de sensores inalámbricos (WSN, Wireless Sensor Networks).

En [3], se propone un sistema EMS basado en simulación mediante CupCarbon, en el cual se integran protocolos de comunicación como WiFi, Zigbee y LoRa para controlar dispositivos como iluminación y aire acondicionado en función de la ocupación del aula. Los resultados muestran una reducción de consumo energético de hasta 51.55% mediante estrategias basadas en ocupación. Asimismo, investigaciones previas han explorado el uso de plataformas como SCADA y herramientas como HOMER para el análisis energético en edificios educativos, así como la implementación de protocolos IoT como MQTT y Node-RED para monitoreo energético en tiempo real [3–5].

## 2.2. IoT aplicado a la eficiencia energética

El IoT ha transformado significativamente la gestión energética al permitir la interconexión de dispositivos y la automatización del consumo eléctrico. Diversos estudios destacan que IoT permite monitorear, controlar y optimizar el uso de energía en edificios mediante sensores distribuidos y comunicación en red [6]. Además, se ha demostrado que la integración de IoT en redes eléctricas inteligentes (smart grids) contribuye a reducir desperdicios energéticos mediante automatización, seguimiento y toma de decisiones en tiempo real. También se han desarrollado sistemas de gestión energética basados en IoT que conectan dispositivos de consumo y producción energética, permitiendo ajustar dinámicamente la demanda y mejorar la eficiencia global del sistema [6,7].

## 2.3 Sensores inteligentes y sistemas basados en contexto

Los sistemas context-aware (sensibles al contexto) han emergido como una evolución de los sistemas IoT tradicionales, incorporando información del entorno para mejorar la toma de decisiones.

Investigaciones previas se han enfocado en sensores de temperatura, humedad y luminancia para controlar dispositivos eléctricos en función de las condiciones ambientales. Por ejemplo, se han desarrollado sistemas de iluminación inteligente que apagan automáticamente las luces cuando no hay ocupación, logrando reducciones significativas en costos y consumo energético.

En [6], reportan que arquitecturas IoT tradicionales presentan ahorros energéticos aproximados del 15% al 20%, mientras que arquitecturas más avanzadas basadas en contexto pueden alcanzar reducciones de entre el 25% y el 30%, además de mejorar el tiempo de procesamiento y ampliar el rango de operación de los sensores.

Sin embargo, muchos de estos enfoques se limitan a datos de sensores físicos, sin considerar el comportamiento humano como variable clave en el consumo energético.

## 2.4 Implementación de arquitecturas IoT de bajo costo para el control de iluminación

Una tendencia creciente en la literatura reciente es el desarrollo de sistemas "Do-It-Yourself" (DIY) que permiten una alta eficiencia sin requerir inversiones elevadas.

En [9], se presenta el sistema ISLS, el cual emplea el microcontrolador ESP32 aprovechando su conectividad Wi-Fi para integrar un ecosistema de sensores de ocupación y luminosidad. A diferencia de soluciones comerciales cerradas, esta arquitectura permite la interacción inalámbrica en tiempo real y el monitoreo remoto de parámetros a través de una aplicación final. El estudio demuestra que el uso de algoritmos de optimización aplicados a hardware de bajo costo puede generar reducciones en el consumo de energía de entre el 26% y el 80%, validando que el control remoto y la automatización basada en microcontroladores estándar son suficientes para lograr objetivos de sostenibilidad en edificios de oficinas y entornos educativos [8—10].

## 2.5 Prototipos de aulas inteligentes basados en IoT para monitoreo ambiental

En [11], se presenta el desarrollo de un prototipo de aula inteligente basado en IoT, el cual utiliza el microcontrolador NodeMCU ESP8266 y la plataforma Blynk para el monitoreo ambiental en tiempo real. Este sistema integra diversos sensores, como el DHT11 para medir temperatura y humedad, el sensor MQ-2 para la detección de gases y humo, y sensores magnéticos para el monitoreo de la apertura y cierre de puertas.

El sistema permite la automatización de dispositivos eléctricos como ventiladores e iluminación mediante módulos de relé, en función de las condiciones ambientales detectadas y la actividad dentro del aula. Asimismo, los datos capturados por los sensores son enviados a una aplicación móvil a través de Blynk, lo que facilita el monitoreo remoto y la toma de decisiones en tiempo real.

Los resultados obtenidos demuestran que el sistema es capaz de operar con una frecuencia de actualización aproximada de un segundo, logrando una respuesta rápida en la activación de dispositivos (menor a un segundo) y una adecuada precisión en los sensores utilizados. Además, se destaca que el prototipo representa una solución de bajo costo y escalable, adecuada para su implementación en entornos educativos.

No obstante, el estudio también identifica algunas limitaciones, como la dependencia de la conectividad WiFi, posibles retrasos en la transmisión de datos y la necesidad de calibración en sensores como el MQ-2 para garantizar mediciones confiables. En general, este trabajo evidencia la importancia de integrar múltiples variables ambientales y de seguridad dentro de un mismo sistema IoT, contribuyendo tanto a la eficiencia energética como a la mejora de las condiciones del entorno educativo [11].

## 2.6 Arquitecturas IoT para automatización energética

Las arquitecturas IoT utilizadas en sistemas de gestión energética suelen basarse en modelos distribuidos compuestos por sensores, actuadores, controladores y plataformas de comunicación. Estas arquitecturas permiten la adquisición, procesamiento y transmisión de datos en tiempo real, facilitando la toma de decisiones automatizada.

Diversos estudios destacan que los sistemas IoT se estructuran en capas, incluyendo una capa de percepción, una capa de red y una capa de aplicación, lo que permite una integración flexible y escalable de dispositivos. Además, el uso de microcontroladores como el ESP32 permite implementar arquitecturas híbridas que combinan procesamiento local y servicios en la nube, mejorando la eficiencia energética y reduciendo la latencia en la respuesta del sistema [12—16].

### 3. Metodología

En esta sección se describe la metodología empleada para el desarrollo del sistema de automatización orientado al ahorro energético en aulas. El proceso se estructuró en diversas etapas, incluyendo el diseño de la arquitectura del sistema, la simulación del circuito, la implementación en el microcontrolador ESP32, la construcción de una maqueta a escala funcional y la integración con una aplicación web para el monitoreo y control remoto. Cada etapa permitió la validación progresiva del sistema propuesto.

La maqueta desarrollada presenta dimensiones aproximadas de  $40 \times 20 \times 30$  cm, representando un aula real de aproximadamente  $5 \times 4 \times 3$  m, lo que corresponde a una escala geométrica no uniforme en el rango de 1:10 a 1:20 dependiendo del eje. Sin embargo, dado que el objetivo principal es la optimización del consumo energético, el escalamiento se aborda desde un enfoque funcional más que estrictamente físico. En este sentido, los actuadores y dispositivos empleados no replican el consumo energético absoluto de un aula real, sino que modelan el comportamiento relativo del sistema bajo distintas condiciones de ocupación y luminosidad.

Este enfoque metodológico se fundamenta en sistemas de control basados en la detección de ocupación, los cuales han demostrado ser efectivos en la reducción del consumo energético. Estudios previos reportan ahorros de hasta un 43.5% en sistemas de iluminación inteligente basados en IoT en comparación con esquemas de operación continua, lo que respalda la validez del enfoque adoptado en este trabajo [17].

#### 3.1. Arquitectura del sistema y diagrama de funcionamiento

El diagrama de funcionamiento fue diseñado con el objetivo de esclarecer los componentes electrónicos a utilizar en el sistema, para así establecer la relación que existirá entre ellos.

La Figura 1 muestra la arquitectura general del sistema propuesto, el cual se basa en el uso del microcontrolador ESP32 como unidad central de procesamiento. El sistema opera bajo dos modos principales: automático y manual. En el modo automático, el sistema recibe información de sensores ambientales, específicamente un sensor ultrasónico para la detección de presencia mediante medición de distancia, y un sensor de humo/gas para la identificación de posibles riesgos en el entorno. Estos datos son procesados por el ESP32, permitiendo la toma de decisiones en tiempo real.

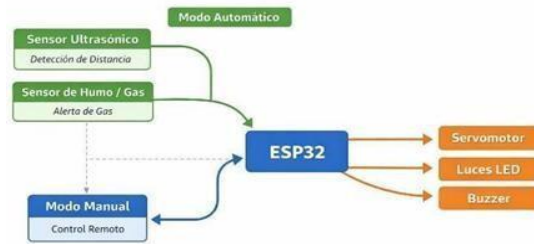


Figura 1. Diagrama de funcionamiento del sistema propuesto. (Fuente: Elaboración propia).

Por otro lado, en el modo manual, el sistema es controlado a través de una aplicación web, la cual permite al usuario interactuar remotamente con los dispositivos conectados, enviando comandos directamente al microcontrolador. Como salida, el sistema controla diversos actuadores, entre los que se incluyen un servomotor, luces LED y un buzzer, los cuales representan, respectivamente, mecanismos de control físico como la apertura y cierre de puertas, sistemas de iluminación del aula y dispositivos de alerta sonora utilizados para notificaciones o advertencias ante condiciones específicas del entorno.

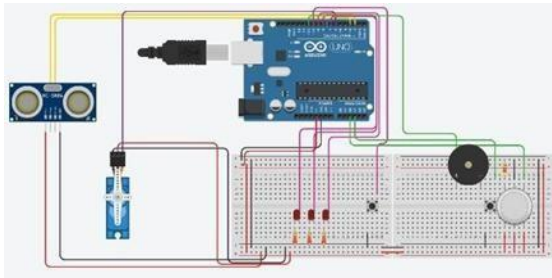
Esta arquitectura híbrida permite combinar automatización basada en sensores con control remoto, proporcionando una solución flexible y eficiente para la gestión energética en aulas.

#### 3.2. Simulación del sistema en Tinkercad

Con el objetivo de validar el funcionamiento del sistema previo a su implementación física, se llevó a cabo la simulación del circuito electrónico mediante la plataforma Tinkercad, desarrollada por Autodesk, la cual permite diseñar, simular y programar circuitos electrónicos en un entorno virtual interactivo sin necesidad de hardware físico. Esta herramienta permitió modelar la interacción entre los distintos componentes

del sistema y verificar su correcto comportamiento en un entorno virtual controlado.

Como se muestra en la Figura 2, la simulación integra un conjunto de dispositivos electrónicos que representan los elementos principales del sistema propuesto. Entre ellos se incluye un sensor ultrasónico para la detección de presencia mediante medición de distancia, un sensor de gas para la identificación de posibles riesgos en el ambiente, un servomotor como actuador mecánico, así como LEDs indicadores y un buzzer para la señalización visual y auditiva. Todos estos componentes se encuentran interconectados y controlados mediante un microcontrolador.



**Figura 2.** Simulación del circuito del sistema en Tinkercad. (Fuente: Elaboración propia).

Es importante destacar que, debido a las limitaciones de la plataforma Tinkercad, la simulación se realizó utilizando un microcontrolador Arduino UNO. Sin embargo, la lógica de control implementada es equivalente a la que posteriormente se desarrolló en el ESP32, manteniendo la misma estructura funcional del sistema.

La simulación permitió verificar la correcta conexión de los componentes, así como la respuesta del sistema ante diferentes condiciones de entrada, tales como la detección de objetos o la presencia de gases. Asimismo, se validó la activación de los actuadores en función de los datos proporcionados por los sensores, garantizando el cumplimiento de la lógica de operación definida en el diseño del sistema.

Estos resultados preliminares facilitaron la identificación y corrección de posibles errores antes de la implementación física, reduciendo riesgos y optimizando el proceso de desarrollo del prototipo.

### 3.3 Implementación del sistema en ESP32

Para la implementación del sistema, se utilizó el microcontrolador ESP32 como unidad central de procesamiento, debido a su capacidad de conectividad inalámbrica y procesamiento en aplicaciones de IoT.

Previo a la conexión física de los componentes, se realizó un análisis del pinout del ESP32 con el objetivo de identificar la distribución y funcionalidad de sus pines, asegurando una correcta asignación de entradas y salidas. Esta etapa fue fundamental para evitar conexiones incorrectas que pudieran dañar el microcontrolador o afectar el funcionamiento del sistema.

A partir del diseño previamente validado en la simulación, se definió la correspondencia entre los pines del ESP32 y los diferentes dispositivos del sistema, incluyendo sensores y actuadores. Este proceso permitió trasladar de manera segura la lógica del circuito simulado hacia una implementación real. Asimismo, se consideraron aspectos como la alimentación de los dispositivos, los niveles de voltaje y la compatibilidad entre componentes, garantizando la integridad del sistema durante su operación.

Esta fase representó un paso clave en la transición de la simulación a la implementación física, permitiendo establecer una base sólida para el desarrollo del prototipo funcional.

### 3.4 Desarrollo de la maqueta a escala funcional

En esta etapa se desarrolló una maqueta con la finalidad de representar el entorno real de aplicación del sistema propuesto. Esta fue construida utilizando materiales ligeros, en su mayoría reciclados, lo que permitió reducir costos sin afectar la funcionalidad del prototipo. La maqueta permite visualizar la integración de los componentes electrónicos dentro de un espacio físico, facilitando la validación del sistema en condiciones representativas.

Como se presenta en la Figura 3 la maqueta presenta una escala geométrica no uniforme respecto a un aula real ( $40 \times 20 \times 30$  cm frente a  $5 \times 4 \times 3$  m). No obstante, el diseño se abordó desde un enfoque funcional, priorizando la interacción entre los dispositivos y el entorno. La distribución interna incluye elementos representativos como mobiliario y estructura del aula, lo que permite simular condiciones de operación cercanas a un escenario real.

La colocación de los sensores se determinó con base en criterios de cobertura espacial y eficiencia en la detección. En particular, los sensores de movimiento se ubicaron en posiciones estratégicas que permiten maximizar el área de detección y minimizar zonas sin cobertura, considerando el ángulo de apertura y el alcance característico de estos dispositivos. Asimismo, se priorizó su instalación en puntos clave del aula, como accesos y zonas centrales, donde se concentra la mayor actividad de los usuarios. Esta disposición permite una detección más confiable de la ocupación del espacio y mejora la respuesta del sistema ante cambios en las condiciones de uso.



**Figura 3.** Maqueta a escala funcional del sistema propuesto. (Fuente: Elaboración propia).

Posteriormente, la maqueta permitió evaluar la ubicación estratégica de los sensores para la detección de variables del entorno, así como la interacción entre los dispositivos y el espacio físico. Esta representación facilitó la observación directa del comportamiento del sistema, permitiendo identificar mejoras en la distribución de componentes y en la lógica de operación.

El desarrollo de la maqueta a escala funcional constituyó una etapa fundamental en la validación del

sistema, ya que permitió comprobar su viabilidad en un entorno físico controlado, previo a su posible implementación en un aula real.

### 3.5 Integración del sistema físico con la aplicación web

Una vez validado el funcionamiento del prototipo físico, se llevó a cabo la integración del sistema con una aplicación web, fue diseñada para proporcionar una interfaz intuitiva que permite al usuario interactuar con el sistema de manera remota. Como se muestra en la Figura 4, la interfaz permite supervisar el estado de los dispositivos y controlar su funcionamiento en tiempo real.

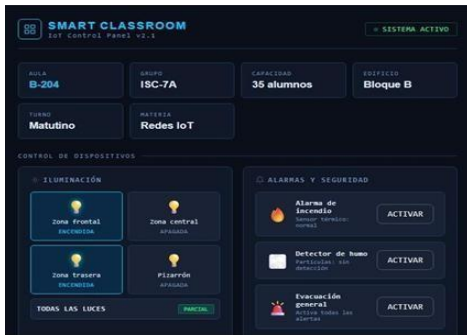


Figura 4. Interfaz web para el control y monitoreo del sistema. (Fuente: Elaboración propia).

La comunicación entre el microcontrolador ESP32 y la aplicación web se estableció mediante una conexión WiFi, basada en el estándar IEEE 802.11 b/g/n en la banda de 2.4 GHz, utilizando el protocolo HTTP para el intercambio de información. A través de este mecanismo, el ESP32 actúa como un servidor capaz de recibir solicitudes provenientes de la interfaz web y responder a ellas mediante la ejecución de acciones específicas en el sistema.

Para estructurar la comunicación, se empleó el formato JSON como medio de intercambio de datos, permitiendo representar de manera organizada las instrucciones enviadas desde la aplicación web hacia el microcontrolador. Mediante el uso de este archivo, se definieron comandos asociados a las distintas funciones del sistema, tales como la activación de actuadores, la lectura de sensores y el cambio entre modos de operación.

Asimismo, fue necesario realizar modificaciones en el código del ESP32 con el fin de adaptar su funcionamiento a la recepción y procesamiento de solicitudes HTTP. Estas modificaciones incluyeron la implementación de rutinas para la interpretación de datos en formato JSON, así como la asignación de funciones específicas en respuesta a las peticiones recibidas.

La aplicación web fue diseñada para proporcionar una interfaz intuitiva que permite al usuario interactuar con el sistema de manera remota, facilitando la supervisión del estado de los dispositivos y el control de estos en tiempo real. Esta integración permite complementar el funcionamiento autónomo del sistema

con capacidades de control remoto, fortaleciendo su aplicabilidad en entornos reales.

### 3.6 Selección de componentes para implementación en aula real

Con el objetivo de evaluar la viabilidad de implementación del sistema en un entorno real, se seleccionaron los componentes necesarios para su despliegue en un aula estándar, considerando criterios de funcionalidad, compatibilidad eléctrica, disponibilidad y costo.

Los elementos incluyen dispositivos de control, sensores, actuadores y componentes de instalación eléctrica, priorizando tecnologías accesibles sin comprometer capacidades operativas como conectividad, comunicación y detección.

La Tabla 1 presenta las especificaciones técnicas y el costo estimado de los componentes, permitiendo analizar la inversión requerida para su implementación.

Tabla 1. Componentes propuestos para implementación en aula real. (Fuente: Elaboración propia).

Componente	Especificaciones	Cantidad	Costo unitario	Costo total
ESP32	Alimentación 5V (USB), lógica 3.3V, Wi-Fi integrado, consumo 80–260 mA	1	200	200
Sensor PIR	Voltaje 110–220V AC, alcance 6–10 m, corriente hasta 10 <sup>a</sup>	3	100	300
Detector de humo fotoeléctrico	Alimentación 127V o batería, alarma 85–100 dB	1	600	600
Módulo de relé de 4 CANALES	Control 5V, capacidad 10 <sup>a</sup> @ 127V	2	100	200
Foco LED 12W	Voltaje 127V, corriente ~0.1 A	6	100	600
Cable THHN	Cableado para	1	600	600

(alta tensión)	instalación eléctrica (30 m)			
Cable multiconductor (baja tensión)	Cableado de control (15 m)	1	150	150
Cerradura eléctrica + cierrapuertas	Sistema de acceso automatizado 12V	1	1200	1200
			<b>Total</b>	3850

A partir de la selección de componentes, se estima que la implementación del sistema en un aula real requiere una inversión aproximada de entre \$3,850 y \$5,000 MXN, lo cual resulta accesible frente a soluciones comerciales que demandan mayores costos e infraestructura especializada.

Los componentes cumplen con los requerimientos técnicos para operar en condiciones reales, incluyendo compatibilidad con redes eléctricas de 127 V, conectividad Wi-Fi y capacidad de detección y control en tiempo real.

Además, la arquitectura propuesta es modular y escalable, lo que facilita su adaptación a distintos entornos educativos sin incrementos significativos en el costo, posicionándose como una alternativa viable para la optimización energética.

#### 4. Resultados Esperados

A partir de las pruebas realizadas en la maqueta a escala, se establecen los resultados esperados para una implementación del sistema en un entorno real. La validación del prototipo permitió comprobar el correcto funcionamiento de los componentes del sistema, incluyendo sensores, actuadores y la unidad de control, así como la comunicación remota, lo cual sirve como base para estimar su desempeño en aulas educativas reales.

En este sentido, se espera que el sistema contribuya a la reducción del consumo energético mediante la automatización del encendido y apagado de dispositivos eléctricos en función de la detección de presencia y condiciones del entorno. Este comportamiento permitiría evitar el uso innecesario de recursos como la iluminación, tal como se describe en la Tabla 2, donde se especifica la función de cada componente dentro del sistema.

Asimismo, se espera que el sistema mantenga una respuesta eficiente en tiempo real, mediante la correcta interacción entre sensores, microcontrolador y actuadores. En particular, el microcontrolador ESP32 actuará como unidad central de procesamiento, gestionando la activación de dispositivos como focos LED, relés y mecanismos de acceso, de acuerdo con las señales recibidas por los sensores.

Por otro lado, en la Tabla 3 se presentan los

resultados esperados relacionados con el desempeño del sistema y la comunicación, donde se establece que la conexión WiFi permitirá una comunicación estable entre el ESP32 y la aplicación web, garantizando el envío y recepción de datos mediante protocolos como HTTP y el uso de estructuras JSON. Se espera que esta comunicación permita el control remoto de los dispositivos de manera eficiente, con tiempos de respuesta rápidos y sin interrupciones durante su operación.

En conjunto, estos resultados esperados permiten establecer una proyección del comportamiento del sistema en condiciones reales, considerando tanto su funcionamiento físico como su desempeño en términos de comunicación y control.

**Tabla 2.** Resultados esperados basados en escala real. (Fuente: Elaboración propia)

Componente	Función dentro del sistema
<b>ESP32</b>	Actúa como unidad central de control, encargada de procesar los datos de los sensores y ejecutar las acciones correspondientes en los actuadores, además de gestionar la comunicación WiFi con la aplicación web.
<b>Sensor PIR</b>	Detecta la presencia de personas dentro del aula mediante movimiento, permitiendo activar o desactivar automáticamente dispositivos eléctricos como la iluminación.
<b>Detector de humo fotoeléctrico</b>	Identifica la presencia de humo o posibles condiciones de riesgo, activando alertas para prevenir situaciones de peligro dentro del aula.
<b>Módulo de relé de 4 canales</b>	Permite controlar dispositivos eléctricos de mayor potencia, como luces o sistemas conectados a corriente alterna, mediante señales del microcontrolador.
<b>Focos LED</b>	Representan el sistema de iluminación del aula, los cuales se activan o desactivan automáticamente en función de la detección de presencia.
<b>Cableado de alta tensión (THHN)</b>	Se utiliza para la conexión de dispositivos que operan con corriente alterna, como

	iluminación y otros equipos eléctricos.
<b>Cableado de baja tensión (multiconductor)</b>	Permite la interconexión de sensores, actuadores y el microcontrolador dentro del sistema electrónico.
<b>Cerradura eléctrica y cierrapuertas hidráulico</b>	Simulan el control de acceso al aula, permitiendo la apertura y cierre automático como parte de la automatización del entorno.

**Tabla 3.** Resultados esperados del desempeño del sistema y comunicación. (Fuente: Elaboración propia).

Parámetro evaluado	Descripción de la prueba	Resultado obtenido
<b>Conectividad WiFi</b>	Conexión continua durante pruebas (30 min)	Conexión estable sin interrupciones
<b>Comunicación HTTP</b>	Envío de múltiples solicitudes desde la app	Respuestas correctas del sistema
<b>Comunicación JSON</b>	Activación remota de actuadores	Respuesta rápida del sistema
<b>Control remoto</b>	Envío y recepción de datos estructurados	Interpretación correcta de comandos
<b>Operación continua</b>	Activación manual desde interfaz web	Ejecución correcta de acciones
<b>Cableado de alta tensión (THHN)</b>	Ejecución del sistema durante pruebas prolongadas	Funcionamiento estable

## 5. Análisis de los Resultados Esperados

Los resultados obtenidos a partir de las pruebas realizadas en la maqueta a escala permiten validar el funcionamiento del sistema en un entorno controlado, evidenciando su capacidad para automatizar y controlar dispositivos de manera remota. Esta validación técnica confirma que la arquitectura propuesta es funcional y puede ser replicada en escenarios reales.

Con base en los resultados presentados en la Tabla

2, se observa que cada componente del sistema cumple una función específica dentro de la automatización del aula. En particular, los sensores permiten la detección de condiciones del entorno, mientras que los actuadores ejecutan acciones como el encendido de iluminación o la activación de mecanismos de acceso.

Esta integración permite establecer un sistema autónomo capaz de responder a eventos en tiempo real.

En relación con la problemática planteada en la introducción, el sistema se orienta a reducir el uso innecesario de dispositivos eléctricos mediante su activación únicamente cuando se detecta presencia. En un entorno real, este comportamiento permitiría optimizar el consumo energético, especialmente en sistemas de iluminación que suelen permanecer encendidos sin requerirse.

Por otra parte, los resultados mostrados en la Tabla 3 evidencian que el sistema es capaz de mantener una comunicación estable mediante WiFi, permitiendo el intercambio de datos entre el microcontrolador y la aplicación web. La correcta transmisión de solicitudes HTTP y la interpretación de datos en formato JSON garantizan el funcionamiento del control remoto, lo cual fortalece la flexibilidad del sistema al combinar automatización y supervisión manual.

Asimismo, la capacidad del sistema para operar de manera continua sin interrupciones durante las pruebas indica que su implementación en un entorno real es viable desde el punto de vista técnico. Sin embargo, es importante considerar aspectos como la dependencia de la conectividad inalámbrica y la calibración de sensores, los cuales pueden influir en el desempeño del sistema en condiciones reales.

En general, los resultados analizados permiten concluir que el sistema propuesto representa una solución funcional y escalable para la gestión energética en aulas, al integrar automatización, monitoreo y control remoto dentro de una misma arquitectura.

## 6. Conclusiones

El presente trabajo permitió diseñar y validar un sistema de automatización de aulas orientado al ahorro energético mediante el uso de tecnologías IoT de bajo costo. A través de la implementación de un prototipo basado en el microcontrolador ESP32 y la integración de sensores y actuadores, fue posible evaluar su funcionamiento en un entorno controlado mediante una maqueta a escala.

Los resultados obtenidos demuestran que el sistema es capaz de detectar condiciones del entorno y actuar en consecuencia, permitiendo el control automático de dispositivos eléctricos como la iluminación y mecanismos de acceso. Asimismo, la integración con una aplicación web permitió validar la viabilidad del control remoto, garantizando una comunicación estable y una respuesta eficiente en tiempo real.

De acuerdo con los resultados esperados y su análisis, el sistema propuesto contribuye a la reducción del consumo energético al evitar el uso innecesario de

recursos, representando una alternativa viable para mejorar la gestión energética en aulas educativas. Además, su arquitectura modular facilita su adaptación a distintos entornos sin requerir inversiones elevadas.

Como trabajo futuro, se plantea la implementación del sistema en un entorno real, así como la incorporación de sensores adicionales que permitan mejorar la precisión del sistema y ampliar sus capacidades de monitoreo.

Finalmente, este proyecto aporta una solución tecnológica accesible para la automatización de espacios educativos, promoviendo el uso eficiente de la energía y el aprovechamiento de tecnologías IoT en instituciones académicas.

## 7. Referencias

- [1] “Gestión inteligente de la energía en instituciones de educación superior | Siemens,” Siemens, 2026. <https://www.siemens.com/es-mx/industries/higher-education/smart-energy-management/> (accessed May 01, 2026).
- [2] “Eficiencia energética de los campus universitarios de la UNED con tecnología IoT e Inteligencia Artificial - Cellnex,” Cellnex, 2022. <https://www.cellnex.com/es-es/casos-uso/eficiencia-energetica-campus-universitarios-uned-iot> (accessed May 01, 2026).
- [3] I. B. Sulistiawati, A. Soetedjo, D. B. Susanti, and G. Sukowiyono, “An Energy Management System in Smart Classroom Using CupCarbon Simulator,” *Engineering, Technology & Applied Science Research*, vol. 15, no. 5, pp. 26985–26995, 2025, doi: 10.48084/etasr.11773.
- [4] L. E. Alarcón Yáñez, E. M. Toscano Díaz, and a. M. E. Oña Llumiquinga, “Prototipo de Gestión Energética en Aulas Domóticas con Android y Arduino Ante el Cambio Climático y Cortes Eléctricos,” *Ibero Ciencias - Revista Científica y Académica - ISSN 3072-7197*, vol. 5, no. 1, pp. 1848–1861, Mar. 2026, doi: <https://doi.org/10.63371/ic.v5.n1.a780>.
- [5] H. A. Sánchez Sánchez, L. Alvarado Tarango, H. S. . Sánchez Rangel, y V. I. . Barrón Ortega, «Diseño del sistema de gestión inteligente para eficientar el uso de la energía a través de la automatización de un edificio en una Institución Superior», *CTES*, vol. 11, n.º 22, nov. 2024.
- [6] P. Paudel, S. Kim, S. Park, and K.-H. Choi, “A Context-Aware IoT and Deep-Learning-Based Smart Classroom for Controlling Demand and Supply of Power Load,” *Electronics*, vol. 9, no. 6, p. 1039, Jun. 2020, doi: 10.3390/electronics9061039.
- [7] J. Vidal and E. Alejandro, “Aplicaciones de internet de las cosas (IoT) en la optimización de la eficiencia energética en entornos residenciales.,” *Revista Científica Multidisciplinar G-nerando*, vol. 6, no. 1, Jan. 2025, doi: <https://doi.org/10.60100/rcmg.v6i1.406>.
- [8] F. Peña de Loza and F. J. Ibarra Villegas, “Implementación de tecnologías IoT para la reducción del consumo energético en oficinas inteligentes mediante el control de la iluminación,” *Revista de Ciencias Tecnológicas (RECIT)*, vol. 7, no. 3, e332, 2024, doi: 10.37636/recit.v7n3e332.
- [9] P. Obioma, O. Agbodike, J. Chen, and L. Wang, “ISLS: IoT- Based Smart Lighting System for Improving Energy Conservation in Office Buildings,” *arXiv preprint arXiv:2503.13474v2*, 2025.
- [10] E. A. Alvarado Bastidas, “Evaluación de sistemas electrónicos de bajo consumo energético en IoT: Aplicaciones educativas para reducir el consumo energético,” *Sage Sphere International Journal*, vol. 1, no. 2, 2024.
- [11] A. Zilni and I. S. Wibisono, “IoT-Based Smart Classroom Prototype Using NodeMCU and Blynk for Environmental Monitoring,” *Journal of Information Systems and Informatics*, vol. 7, no. 4, Dec. 2025, doi: 10.63158/journalisi.v7i4.1356.
- [12] S. Bhaganagare, S. Chavan, S. Gavali y V. V. Godase, “Voice-Controlled Home Automation with ESP32: A Systematic Review of IoT- Based Solutions,” *Journal of Microprocessor and Microcontroller Research*, vol. 2, no. 3, pp. 1–13, 2025.
- [13] C. J. Amaya Rondón, “Evaluación de la automatización para el ahorro energético en edificios educativos,” Tesis de maestría, Universidad de la Costa, 2024.
- [14] E. A. Morocho-Belduma, B. S. Masache- Carrera, D. Hernandez-Rojas, y B. Mazon- Olivo, «Optimización energética en entornos académicos mediante IoT y aprendizaje automático», *I.P.*, vol. 8, n.º 14, pp. 4–26, ene. 2025.
- [15] R. Javier, “Sistema internet de las cosas IOT, para la automatización y monitoreo de iluminación y climatización: impulso a la eficiencia energética en aulas universitarias,” *Upse.edu.ec*, Feb. 04, 2026. <https://repositorio.upse.edu.ec/handle/46000/16268> (accessed Apr. 09, 2026).
- [16] R. Enrique, “Diseño e Implementación de un Módulo de entrenamiento utilizando el procesador ESP32 para aplicaciones enfocadas a la Domótica,” *Ups.edu.ec*, 2022, doi: <http://dspace.ups.edu.ec/handle/123456789/22860>.
- [17] S. S. M. Y. Grandhi, “An IoT-Based Smart Lighting System for Real-Time Occupancy Monitoring and Energy Management,” *European Journal of Advances in Engineering and Technology*, vol. 9, no. 5, pp. 174–182, 2022.

# Propuesta de un Visor Geográfico en la ciudad Mazatlán, Sinaloa

## Proposal for a Geographic Viewer in the city of Mazatlán, Sinaloa

Michelle Quintero-Bonilla<sup>1</sup>, Yaritza Alejandra Laureano-delaCruz<sup>1</sup>, Luis Jesús Vázquez-Osuna<sup>1</sup>, Zeus Del Valle-Castillo Nájera<sup>2</sup>

<sup>1</sup>Facultad de Informática Mazatlán, Universidad Autónoma de Sinaloa, México.

<sup>2</sup>Facultad de Informática Culiacán, Universidad Autónoma de Sinaloa, México.

Michelle Quintero Bonilla, qbmichelle175@gmail.com, ORCID:0009-0005-2031-6744

Yaritza Alejandra Laureano de la Cruz, eyaritza90@gmail.com, ORCID: 0009-0001-9166-7999

Luis Jesus Vazquez Osuna, lvguijesus@gmail.com, ORCID: 0009-0007-8344-8933

Autor por correspondencia: Zeus Del Valle-Castillo Nájera, zeus@uas.edu.mx, ORCID: <https://orcid.org/0009-0000-8265-6035>

**Recibido:** abril 2026, **Aceptado:** abril 2026, **Publicado:** mayo 2026

### Resumen:

El crecimiento urbano acelerado en ciudades intermedias como Mazatlán, Sinaloa, genera una demanda creciente de herramientas tecnológicas para la gestión y visualización de información territorial. En este contexto, el presente trabajo propone el diseño arquitectónico de un visor geográfico web orientado a centralizar y facilitar el acceso a datos geoespaciales del municipio mediante una plataforma interactiva de código abierto. La arquitectura propuesta integra Leaflet.js como motor de visualización cartográfica en el cliente, PostgreSQL con la extensión PostGIS para la gestión de datos espaciales en el servidor, y servicios estándar del Open Geospatial Consortium (OGC), como WMS y WFS, para garantizar la interoperabilidad entre capas de información. La metodología comprende cuatro etapas: revisión sistemática de la literatura, recopilación y estructuración de datos geoespaciales institucionales, diseño de la arquitectura del sistema y validación conceptual mediante diagramas de flujo y casos de uso. Se espera que el sistema permite integrar capas temáticas de infraestructura urbana, delimitaciones territoriales y puntos de interés del municipio, haciéndolas accesibles a ciudadanos, investigadores y tomadores de decisiones sin necesidad de software especializado. Este trabajo contribuye a la democratización de la información geoespacial en el ámbito municipal y sienta las bases técnicas para la implementación futura del sistema.

### Palabras Clave:

*Sistemas de Información Geográfica (SIG), WebGIS, visor geográfico web, análisis espacial, planificación urbana, infraestructura de datos geoespaciales, SIG de código abierto, gestión territorial, ciudades inteligentes, visualización de datos espaciales, morfología urbana, crecimiento urbano, ordenamiento territorial, datos abiertos geoespaciales.*

### Abstract:

Accelerated urban growth in intermediate cities such as Mazatlán, Sinaloa, has generated an increasing demand for technological tools for the management and visualization of territorial information. In this context, this paper proposes the architectural design of a web-based geographic viewer aimed at centralizing and facilitating access to the municipality's geospatial data through an open-source interactive platform. The proposed architecture integrates Leaflet.js as the client-side cartographic visualization engine, PostgreSQL with the PostGIS extension for server-side spatial data management, and standard services from the Open Geospatial Consortium (OGC), such as WMS and WFS, to ensure interoperability between data layers. The methodology comprises four stages: systematic literature review, collection and structuring of institutional geospatial data, system architecture design, and conceptual validation through flow diagrams and use cases. The system is expected to integrate thematic layers of urban infrastructure, territorial boundaries, and points of interest, making them accessible to citizens, researchers, and decision-makers without the need for specialized software. This work contributes to the democratization of geospatial information at the municipal level and establishes the technical foundation for the future implementation of the system.

---

**Keywords:**

*Geographic Information Systems (GIS), WebGIS, web-based geographic viewer, spatial analysis, urban planning, geospatial data infrastructure, open-source GIS, territorial management, smart cities, spatial data visualization, urban morphology, urban growth, land-use planning, open geospatial data.*

## 1. Introducción

El crecimiento demográfico y la expansión territorial de las ciudades latinoamericanas han incrementado la complejidad de la gestión urbana, haciendo indispensable el uso de herramientas que permitan analizar y visualizar datos geoespaciales de manera eficiente y accesible. Los Sistemas de Información Geográfica (SIG) constituyen una tecnología fundamental en este proceso, al integrar la captura, almacenamiento, análisis y visualización de información georreferenciada para apoyar la toma de decisiones en múltiples ámbitos: planificación territorial, gestión de infraestructura, análisis de riesgos y participación ciudadana [2,3].

Desde una perspectiva urbanística, el crecimiento de las ciudades intermedias en América Latina ha sido caracterizado por procesos de expansión dispersa, fragmentación territorial y transformaciones en la morfología urbana, lo que dificulta la planificación eficiente del uso del suelo y la provisión de servicios urbanos. Estas dinámicas generan patrones complejos de ocupación del territorio que requieren herramientas de análisis espacial capaces de integrar múltiples variables geográficas. En este sentido, los SIG permiten comprender fenómenos como la densificación, la segregación socioespacial y la distribución de equipamientos urbanos, constituyéndose como un soporte clave para la gestión urbana contemporánea.

En los últimos años, el desarrollo de visores geográficos web ha democratizado el acceso a la información espacial al eliminar la necesidad de instalar software especializado en el equipo del usuario. Estas plataformas, construidas sobre tecnologías de código abierto como Leaflet.js, GeoServer y PostGIS, permiten consultar capas de información territorial desde cualquier dispositivo con acceso a internet, facilitando la interacción con mapas interactivos, la superposición de datos temáticos y la realización de consultas espaciales básicas [4,5].

En el contexto de Mazatlán, Sinaloa, ciudad portuaria con una superficie municipal de 3,068 km<sup>2</sup> y una población superior a 500,000 habitantes según el Censo de Población y Vivienda 2020 [6], la gestión del territorio adquiere especial relevancia debido a su acelerado crecimiento urbano, impulsado por el desarrollo turístico, la actividad portuaria y la expansión de zonas residenciales. Este crecimiento genera una demanda creciente de información geoespacial actualizada y accesible para autoridades, planificadores urbanos y ciudadanía en general.

No obstante, la información territorial disponible en Mazatlán, proveniente de instituciones como el Instituto Municipal de Planeación (IMPLAN), el INEGI y dependencias estatales, se encuentra dispersa y, en muchos casos, no está integrada en plataformas de acceso

público que permitan su consulta de manera unificada. Esta fragmentación limita el aprovechamiento del acervo geoespacial existente y dificulta la toma de decisiones basada en evidencia en procesos de planificación urbana.

Ante esta problemática, el presente trabajo propone el diseño arquitectónico de un visor geográfico web para la ciudad de Mazatlán, fundamentado en tecnologías de código abierto ampliamente adoptadas en la comunidad científica [4,5], [8]–[10]. La propuesta busca centralizar, organizar y visualizar información geoespacial relevante en una plataforma accesible, contribuyendo así a la modernización de la gestión territorial municipal y al fortalecimiento de infraestructuras de datos espaciales a nivel local.

El presente artículo se estructura de la siguiente manera: en la sección 2 se presentan los trabajos relacionados y fundamentos teóricos; en la sección 3 se describe la metodología empleada para el diseño del visor geográfico; en la sección 4 se exponen los resultados esperados y las funcionalidades del sistema; en la sección 5 se realiza el análisis de la propuesta; y finalmente, en la sección 6 se presentan las conclusiones y las líneas de trabajo futuro.

## 2. Trabajos Relacionados

La adopción de visores geográficos web basados en tecnologías de código abierto ha sido documentada en diversas investigaciones [4]. En este contexto, se desarrolló Farm GeoBalance, un sistema WebGIS para la gestión de prácticas agrícolas en Grecia que combina Leaflet.js, PostGIS y PHP. Los autores reportaron que la integración de una geodatabase espacial con una interfaz web interactiva redujo significativamente los tiempos de consulta de información territorial, demostrando la viabilidad de esta arquitectura en contextos de gestión aplicada.

En el ámbito de la salud pública, Mileva et al. [5] implementaron una aplicación Web GIS de código abierto para evaluar la distribución geográfica de indicadores de calidad sanitaria en Portugal, utilizando Leaflet, PostGIS y GeoServer. El estudio evidenció que esta combinación de herramientas permite construir prototipos funcionales con capacidad de análisis espacial sin incurrir en costos de licenciamiento, lo que resulta especialmente relevante para municipios con presupuestos tecnológicos limitados.

Desde la perspectiva teórica, González [2] analizó la evolución de los SIG hacia representaciones geométricas más complejas, subrayando la importancia de los visores web como interfaces de acceso a datos geoespaciales en entornos municipales. Por su parte, Buzai y Montes Galbán [3] examinaron el rol de los SIG en la planificación territorial latinoamericana, destacando la brecha existente entre la disponibilidad de datos geoespaciales institucionales y las herramientas de acceso público a dichos datos.

En el contexto mexicano, aunque existen iniciativas nacionales como los servicios WMS de la Infraestructura de Datos Espaciales de México (IDEM) [11], la implementación de visores geográficos a escala municipal orientados a la consulta pública permanece limitada en ciudades intermedias como Mazatlán. El presente trabajo busca atender este vacío mediante una propuesta arquitectónica replicable y de bajo costo.

Estos resultados coinciden con investigaciones recientes sobre plataformas WebGIS y ciudades inteligentes, donde el uso de software geoespacial de código abierto ha demostrado ser escalable e interoperable [8]-[10].

## 2.1 Acrónimos

Para garantizar la claridad terminológica, se establecen las siguientes definiciones:

1. SIG (Sistema de Información Geográfica): sistema tecnológico que integra hardware, software, datos y procedimientos para capturar, almacenar, analizar y visualizar información georeferenciada [2].
2. WMS (Web Map Service): estándar del Open Geospatial Consortium (OGC) para la publicación de imágenes cartográficas a través de internet [1].
3. WFS (Web Feature Service): estándar OGC que permite consultar y modificar datos vectoriales geoespaciales a través de servicios web [1].
4. IDE (Infraestructura de Datos Espaciales): conjunto de políticas, estándares y tecnologías que facilitan la producción, gestión, acceso e intercambio de información geoespacial entre organizaciones [2].

Desde el enfoque de análisis urbano inteligente, Li et al. [16] propusieron una plataforma geoespacial basada en big data para ciudades inteligentes, integrando sensores IoT y análisis espacial en tiempo real. Los autores concluyeron que los visores geográficos constituyen componentes esenciales para la gobernanza urbana basada en datos.

Asimismo, Zhao et al. [17] desarrollaron un marco WebGIS para monitoreo urbano dinámico utilizando computación en la nube y servicios geoespaciales distribuidos, demostrando mejoras importantes en escalabilidad y procesamiento espacial.

En el contexto de infraestructura geoespacial abierta, Brovelli et al. [18] analizaron el uso de aplicaciones SIG participativas para fortalecer procesos de participación ciudadana y recopilación colaborativa de información territorial. Los autores resaltan que las plataformas abiertas permiten reducir barreras tecnológicas y ampliar el acceso a información pública.

Adicionalmente, estudios recientes sobre interoperabilidad geoespacial han mostrado que las

arquitecturas basadas en microservicios y APIs geoespaciales favorecen la integración entre plataformas institucionales distribuidas [19, 20]. Estas tendencias representan un cambio relevante respecto a modelos tradicionales monolíticos de gestión territorial.

En América Latina, Buzai y Montes Galbán [21] identificaron importantes desafíos relacionados con fragmentación de datos geoespaciales, limitada interoperabilidad institucional y ausencia de plataformas municipales abiertas. Los autores enfatizan la necesidad de desarrollar infraestructuras de datos espaciales accesibles para fortalecer procesos de planificación territorial.

A pesar del crecimiento de investigaciones sobre plataformas WebGIS, continúan existiendo limitaciones importantes en ciudades intermedias latinoamericanas, particularmente en relación con la integración municipal de información geoespacial. En el caso específico de Mazatlán, no se identificaron estudios científicos orientados al diseño de un visor geográfico interoperable basado en estándares abiertos y tecnologías WebGIS modernas.

En consecuencia, el presente trabajo contribuye al estado del arte mediante una propuesta arquitectónica orientada a contextos urbanos intermedios, integrando estándares OGC, tecnologías abiertas y principios de interoperabilidad geoespacial para fortalecer procesos de gestión territorial inteligente.

## 3. Metodología

El presente trabajo adopta un enfoque metodológico aplicado, orientado al diseño de la arquitectura de un visor geográfico web para la ciudad de Mazatlán, Sinaloa. La metodología se estructura en cinco fases principales: (1) análisis de requerimientos, (2) identificación y recopilación de datos geoespaciales, (3) diseño de la arquitectura del sistema, (4) selección de tecnologías y (5) definición de funcionalidades del visor.

### 3.1 Análisis de requerimientos

En esta fase se identificaron las necesidades de los potenciales usuarios del sistema, considerando tres perfiles principales: autoridades municipales, planificadores urbanos y ciudadanía en general. El análisis se centró en determinar los tipos de información geoespacial requeridos, así como las funcionalidades necesarias para su consulta y visualización.

Entre los requerimientos funcionales identificados destacan:

- Visualización de capas geográficas temáticas (uso de suelo, infraestructura, equipamiento urbano).
- Navegación interactiva (zoom, desplazamiento y selección de elementos).

- Consulta de atributos asociados a entidades espaciales.
- Superposición de capas de información.
- Acceso desde dispositivos con conexión a internet sin necesidad de software especializado.

Asimismo, se definieron requerimientos no funcionales relacionados con la usabilidad, escalabilidad, interoperabilidad y rendimiento del sistema.

Este análisis de requerimientos se encuentra directamente vinculado con las problemáticas urbanas identificadas en el municipio, tales como la expansión territorial, la distribución desigual de servicios y la necesidad de acceso a información geoespacial integrada. En este sentido, el visor geográfico se plantea como una herramienta que facilita la comprensión de dinámicas urbanas y apoya la toma de decisiones en planificación territorial.

### 3.2 Identificación y recopilación de datos geoespaciales

La información geoespacial considerada en este estudio proviene de diversas fuentes institucionales, entre las que destacan el Instituto Nacional de Estadística y Geografía (INEGI), el Instituto Municipal de Planeación (IMPLAN) de Mazatlán y dependencias gubernamentales estatales.

Los datos recopilados incluyen:

Cartografía base (límites administrativos, red vial, hidrografía).

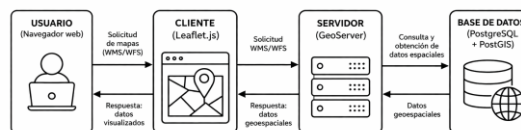
- Información sociodemográfica.
- Uso de suelo y zonificación.
- Infraestructura urbana y equipamiento.

Estos datos fueron organizados y estandarizados para garantizar su compatibilidad e integración dentro del sistema. Asimismo, se consideró el uso de estándares abiertos para la publicación y consulta de servicios geoespaciales, como los definidos por el Open Geospatial Consortium (OGC), incluyendo Web Map Service (WMS) y Web Feature Service (WFS) [1].

Los criterios de selección de los datos geoespaciales se basaron en su relevancia para el análisis urbano del municipio. En particular, las capas de uso de suelo permiten identificar patrones de crecimiento y expansión urbana; la red vial facilita el análisis de accesibilidad y conectividad; la infraestructura y equipamiento urbano contribuyen a evaluar la cobertura de servicios; mientras que la información sociodemográfica permite analizar la distribución espacial de la población. La integración de estas capas resulta fundamental para apoyar procesos de planificación territorial y toma de decisiones basadas en evidencia.

### 3.3 Diseño de la Arquitectura del Sistema

El visor geográfico web se plantea bajo una arquitectura cliente-servidor de tres capas:



**Figura 1.** Arquitectura del visor geográfico propuesto basada en un modelo cliente-servidor de tres capas.

- Capa de datos: encargada del almacenamiento y gestión de la información geoespacial mediante una base de datos espacial.
- Capa de servicios: responsable de la publicación de datos a través de servicios web geoespaciales interoperables.
- Capa de presentación: interfaz web que permite la interacción del usuario con el visor.

Esta arquitectura permite una adecuada separación de responsabilidades, facilitando la escalabilidad, mantenimiento y reutilización de los componentes del sistema.

### 3.4 Selección de tecnologías

La propuesta se basa en el uso de tecnologías de código abierto ampliamente utilizadas en el ámbito de los Sistemas de Información Geográfica. Entre las principales herramientas seleccionadas se encuentran:

- PostgreSQL/PostGIS: para el almacenamiento y gestión de datos espaciales.
- GeoServer: para la publicación de servicios geoespaciales (WMS/WFS).
- Leaflet.js: para la visualización interactiva de mapas en el navegador web.

La elección de estas tecnologías responde a criterios de interoperabilidad, flexibilidad, bajo costo de implementación y amplia documentación disponible.

### 3.5 Proceso metodológico del visor geográfico

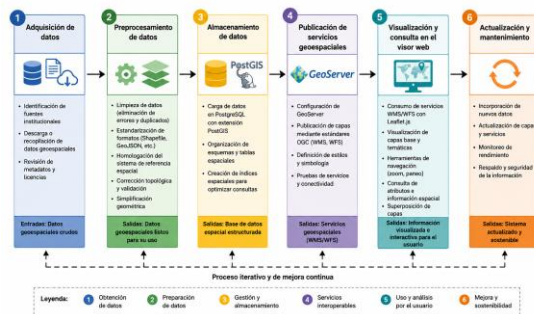


Figura 2. Diagrama del proceso metodológico para el desarrollo del visor geográfico web.

El proceso metodológico se estructura en un flujo secuencial que inicia con la adquisición de datos geoespaciales, seguido del preprocesamiento y estandarización de la información, su almacenamiento en una base de datos espacial, la publicación mediante servicios interoperables y finalmente su visualización en una interfaz web interactiva. Este flujo garantiza la calidad, interoperabilidad y accesibilidad de la información territorial.

Las etapas del proceso son las siguientes:

#### 3.5.1 Adquisición de datos

Consiste en la recopilación de información geoespacial proveniente de diversas fuentes institucionales, tales como el INEGI, IMPLAN Mazatlán y dependencias gubernamentales. Los datos incluyen cartografía base, información sociodemográfica, uso de suelo e infraestructura urbana.

#### 3.5.2 Preprocesamiento de datos

En esta etapa se realiza la preparación de los datos geoespaciales para su integración en el sistema. Incluye:

- Limpieza de datos (eliminación de inconsistencias y duplicados)
- Estandarización de formatos (Shapefile, GeoJSON, etc.)
- Homologación de sistemas de referencia espacial
- Corrección topológica
- Simplificación geométrica para optimizar el rendimiento

Este proceso es fundamental para asegurar la calidad de la información y su correcto funcionamiento dentro del visor.

#### 3.5.3 Almacenamiento de datos

Los datos preprocesados se almacenan en una base de datos espacial utilizando PostgreSQL con la extensión PostGIS, lo que permite gestionar eficientemente grandes volúmenes de información geográfica y realizar consultas espaciales.

#### 3.5.4 Publicación de servicios geoespaciales

Los datos almacenados son publicados mediante servicios interoperables utilizando Geoserver, implementando estándares del Open Geospatial Consortium (OGC), tales como Web Map Service (WMS) y Web Feature Service (WFS)<sup>1</sup>. Esto permite que la información sea accesible desde aplicaciones web y otros sistemas SIG.

#### 3.5.5 Visualización y consulta

Finalmente, los servicios geoespaciales son consumidos por la interfaz web desarrollada con Leaflet.js, donde los usuarios pueden:

- Visualizar capas temáticas
- Navegar en el mapa
- Consultar atributos espaciales
- Superponer información geográfica

#### 3.5.6 Actualización y mantenimiento

El sistema contempla la posibilidad de integrar nuevos datos y actualizar la información existente, garantizando la vigencia del visor geográfico como herramienta de apoyo a la gestión territorial.

### 3.6 Síntesis metodológica

Los resultados del presente trabajo corresponden a una propuesta metodológica orientada al diseño de un visor geográfico web, la cual establece las bases técnicas y conceptuales para su implementación futura. En este sentido, el estudio se enfoca en definir la arquitectura, los componentes del sistema y las funcionalidades esperadas, más que en la evaluación de un sistema ya implementado.

La metodología propuesta integra aspectos técnicos y conceptuales del desarrollo de sistemas geoespaciales, orientándose a la creación de una herramienta accesible y funcional para la gestión territorial. El enfoque adoptado permite no sólo estructurar el desarrollo del visor geográfico, sino también establecer una base replicable para futuras implementaciones en contextos urbanos similares.

### 4. Resultados

Dado que el presente estudio se enfoca en el diseño metodológico y arquitectónico del visor geográfico, los resultados se presentan en términos de la definición de componentes, funcionalidades y estructura del sistema propuesto. En este sentido, el trabajo se enmarca como un artículo de carácter metodológico, orientado a establecer

una base técnica replicable para el desarrollo de plataformas WebGIS en contextos urbanos.

#### 4.1 Capas Temáticas

El visor propuesto integra:

- (a) delimitaciones territoriales;
- (b) infraestructura urbana;
- (c) uso de suelo y zonificación;
- (d) puntos de interés.

#### 4.2 Funcionalidades

El sistema contempla las siguientes funcionalidades:

1. Navegación (zoom y desplazamiento)
2. Cambio de mapa base
3. Consulta de atributos
4. Activación de capas
5. Medición de distancias y áreas
6. Búsqueda por colonia
7. Exportación en GeoJSON

#### 4.3 Métricas de Evaluación

1. Tiempo de respuesta inferior a 2 segundos
2. Evaluación de usabilidad (SUS)
3. Disponibilidad del sistema: 99.5%

#### 4.4 Prototipo funcional

Como parte de la validación de la propuesta, se plantea el desarrollo de un prototipo funcional del visor geográfico que permitirá visualizar capas base, activar información temática y realizar consultas espaciales básicas. Este prototipo constituirá una línea de trabajo futuro para la validación empírica del modelo propuesto, permitiendo evaluar su rendimiento, usabilidad y aplicabilidad en la gestión urbana.

### 5. Análisis de Resultados

La propuesta se alinea con el uso de plataformas Web GIS de código abierto. La arquitectura Leaflet + PostGIS + GeoServer ha demostrado ser escalable en distintos contextos, permitiendo manejar grandes volúmenes de datos, tal como ha sido reportado en estudios recientes sobre sistemas WebGIS, ciudades inteligentes y plataformas geoespaciales abiertas [8]-[10].

Una ventaja frente a soluciones comerciales es la eliminación de costos de licenciamiento. Además, el uso de estándares OGC permite la interoperabilidad con otras plataformas.

Como limitación, al tratarse de un diseño conceptual, el desempeño deberá validarse en implementación. Asimismo, la calidad y actualización de los datos geoespaciales puede influir en la utilidad del sistema.

Además, en comparación con los estudios revisados, la propuesta destaca por su enfoque en la integración de datos geoespaciales a nivel municipal, lo cual representa una contribución relevante para contextos urbanos intermedios donde la información suele estar fragmentada. Esto permite no solo mejorar la accesibilidad a los datos, sino también fortalecer los procesos de planificación territorial basados en evidencia.

La arquitectura propuesta presenta ventajas significativas respecto a modelos tradicionales de gestión territorial basados en sistemas SIG de escritorio aislados. En primer lugar, el uso de tecnologías abiertas reduce considerablemente los costos de implementación y mantenimiento, aspecto crítico para municipios con restricciones presupuestales.

En segundo término, la adopción de estándares OGC garantiza interoperabilidad entre diferentes plataformas geoespaciales, permitiendo integrar información proveniente de múltiples dependencias institucionales. Esta característica resulta particularmente relevante en contextos urbanos donde los datos territoriales suelen encontrarse fragmentados entre diversas entidades gubernamentales.

Desde la perspectiva tecnológica, la combinación PostgreSQL/PostGIS + GeoServer + Leaflet.js representa una arquitectura ampliamente validada en aplicaciones científicas recientes debido a su escalabilidad, estabilidad y compatibilidad con infraestructuras geoespaciales modernas.

Asimismo, la propuesta se alinea con tendencias contemporáneas asociadas con ciudades inteligentes, gobernanza digital y análisis urbano basado en datos. La futura incorporación de sensores IoT, analítica espacial en tiempo real y participación ciudadana permitiría evolucionar el sistema hacia una plataforma geoespacial inteligente orientada a la gestión urbana integral.

No obstante, el estudio presenta limitaciones inherentes a su naturaleza metodológica y conceptual. La validación empírica del rendimiento, usabilidad y escalabilidad deberá realizarse durante la etapa futura de implementación funcional. Igualmente, la calidad y actualización permanente de los datos geoespaciales constituye un factor crítico para garantizar la utilidad operativa del visor.

### 6. Conclusiones

El presente trabajo propone el diseño arquitectónico de un visor geográfico web para la ciudad de Mazatlán, Sinaloa. Esta propuesta se fundamenta en tecnologías de código abierto como Leaflet.js, PostGIS y GeoServer, ampliamente documentadas en la literatura científica reciente por su flexibilidad, escalabilidad e interoperabilidad en sistemas WebGIS [8]-[15], así como en estándares del Open Geospatial Consortium (OGC), como el servicio Web Map Service (WMS) [1].

Asimismo, atiende la necesidad de centralizar y democratizar el acceso a información territorial actualmente dispersa entre diversas instituciones municipales y estatales.

Las principales contribuciones de este trabajo son:

1. Definición de una arquitectura técnica replicable en contextos municipales similares, con bajo costo de implementación.
2. Identificación y categorización de las fuentes de datos geospaciales institucionales disponibles para el municipio de Mazatlán.
3. Establecimiento de métricas de evaluación concretas para orientar la fase de implementación futura.

Como trabajo futuro, se plantea el desarrollo del prototipo funcional del sistema, su evaluación empírica con los perfiles de usuario identificados, y la incorporación de funcionalidades avanzadas como análisis de accesibilidad urbana, visualización de datos en tiempo real provenientes de sensores IoT, y módulos de participación ciudadana para la validación colaborativa de información territorial.

Este trabajo proporciona una base metodológica y técnica que puede ser replicada en otros municipios con características similares, contribuyendo al fortalecimiento de infraestructuras de datos espaciales a nivel local.

Los resultados del estudio evidencian que las tecnologías abiertas representan una alternativa viable para fortalecer infraestructuras municipales de datos espaciales, particularmente en ciudades latinoamericanas con limitaciones presupuestales y fragmentación institucional.

## 7. Referencias

- [1] Open Geospatial Consortium (OGC), "OGC Web Map Service (WMS) Implementation Specification, Version 1.3.0," 2022. [Online]. Available: <https://www.ogc.org/standard/wms/>
- [2] J. A. G. González, "De la topología a la geometría: implementación de mapas mentales a los Sistemas de Información Geográfica," *Cuadernos Geográficos*, vol. 61, no. 2, pp. 88–107, 2022, DOI: 10.30827/cuadgeo.v61i2.22859.
- [3] G. D. Buzai y E. Montes Galbán, "Sistemas de Información Geográfica en la planificación territorial latinoamericana: avances y perspectivas," *Cuadernos Geográficos*, vol. 60, no. 1, pp. 6–27, 2021, DOI: 10.30827/cuadgeo.v60i1.13829.
- [4] T. Mavridis, A. Anastasiou, I. Stavros y I. Kalogeropoulos, "Development of a multi-scale interactive web-GIS system to monitor farming practices: A case study in Lemnos Island, Greece," *Smart Agricultural Technology*, vol. 5, Art. no. 100295, 2023, DOI: 10.1016/j.atech.2023.100295.
- [5] N. Mileva, A. I. Santana y C. M. Lopes, "An Open Source GIS Application for Spatial Assessment of Health Care Quality Indicators," *ISPRS International Journal of Geo-Information*, vol. 10, no. 4, p. 264, 2021, DOI: 10.3390/ijgi10040264.
- [6] INEGI, "Censo de Población y Vivienda 2020: Resultados por municipio," Instituto Nacional de Estadística y Geografía, 2020. [Online]. Available: <https://www.inegi.org.mx/programas/ccpv/2020/>
- [7] IMPLAN Mazatlán, "Plan de Desarrollo Urbano Municipal de Mazatlán 2023–2040," Instituto Municipal de Planeación de Mazatlán, 2023. [Online]. Available: <https://implanmazatlan.gob.mx>
- [8] A. Brovelli, M. Minghini y G. Zamboni, "Public participation in GIS via mobile applications," *ISPRS International Journal of Geo-Information\**, vol. 7, no. 5, 2018, DOI: 10.3390/ijgi7050199.
- [9] F. P. García, J. L. Ariza y M. A. Ureña, "Web GIS for urban planning: A case study," *Computers, Environment and Urban Systems*, vol. 70, pp. 112–123, 2018.
- [10] F. J. Aguilar, M. Aguilar y F. Agüera, "Web-based GIS for urban planning and smart cities," *ISPRS International Journal of Geo-Information*, 2020. DOI: 10.3390/ijgi9090545.
- [11] E. Quirós y M. E. Polo, "Recursos abiertos de información geográfica para investigación y documentación científica," *Revista Española de Documentación Científica*, vol. 41, no. 3, 2018, DOI: 10.3989/redc.2018.3.1512.
- [12] B. Kitchenham y S. Charters, "Guidelines for performing Systematic Literature Reviews in Software Engineering," Keele University, 2007.
- [13] M. Okoli, "A Guide to Conducting a Systematic Literature Review of Information Systems Research," *Sprouts: Working Papers on Information Systems*, vol. 10, no. 26, 2010.
- [14] S. Chowdhury, D. Schröder y H. Ostadabbas, "Modernizing geospatial services: OGC API vs traditional standards," *ISPRS Archives*, 2024. DOI: 10.5194/isprs-archives-XLVIII-4-W12-2024-11-2024.
- [15] P. Campagna, "GIS for sustainable development: open-source web platforms," *Applied Geomatics*, 2021. DOI: 10.1007/s12518-020-00336-3.
- [16] J. Li, H. Zhang, and Y. Wu, "Big geospatial data analytics framework for smart city applications," *Sustainable Cities and Society*, vol. 89, pp. 104118, 2023.
- [17] Y. Zhao, X. Liu, and T. Wang, "Distributed cloud-WebGIS framework for dynamic urban monitoring," *Future Generation Computer Systems*, vol. 141, pp. 20–35, 2023.
- [18] A. Brovelli, M. Minghini, and G. Zamboni, "Citizen participation and collaborative mapping using open-source GIS platforms," *ISPRS International Journal of Geo-Information*, vol. 7, no. 5, pp. 1–19, 2018.
- [19] R. Singh and K. Sharma, "Microservice architectures for geospatial interoperability," *Computers, Environment and Urban Systems*, vol. 102, pp. 101945, 2024.
- [20] D. Ferreira, L. Costa, and M. Ribeiro, "Interoperable geospatial APIs for smart governance systems," *Journal of Urban Technology*, vol. 31, no. 2, pp. 55–77, 2024.
- [21] G. D. Buzai and E. Montes Galbán, "Spatial data infrastructures and territorial planning in Latin America," *Cuadernos Geográficos*, vol. 60, no. 1, pp. 6–27, 2021.

# Asistente inteligente para consultas académicas y administrativas de la Facultad de Informática de la UAS

Intelligent assistant for academic and administrative inquiries of the Faculty of Informatics at UAS

Cristhian Alexis Ortiz Valentin<sup>1</sup>, Héctor Manuel Padilla Osuna<sup>1</sup>, Héctor Joaquin Escobar Cuevas<sup>1</sup>, Manuel Camacho Martinez<sup>1</sup>

<sup>1</sup>Facultad de Informática Mazatlán, Universidad Autónoma de Sinaloa, México.

Cristhian Alexis Ortiz Valentin, ortizcristhian503@gmail.com, 0009-0002-1173-4002

Héctor Manuel Padilla Osuna, gamerspy2003@gmail.com, ORCID: 0009-0002-6378-858X

Manuel Camacho Martínez, manuelmcm645@gmail.com, ORCID: 0009-0005-2835-3400

Autor por correspondencia: Héctor Joaquín Escobar Cuevas, hector.escobar@uas.edu.mx, ORCID: 0000-0002-8845-7069

**Recibido:** abril 2026, **Aceptado:** abril 2026, **Publicado:** mayo 2026

## Resumen:

El presente trabajo propone el diseño e implementación de un asistente académico inteligente orientado a la atención de consultas académicas y administrativas en la Facultad de Informática de la Universidad Autónoma de Sinaloa (UAS). Ante el incremento en la demanda de información y la saturación de los canales tradicionales, se plantea una solución basada en inteligencia artificial para automatizar y optimizar la gestión del conocimiento institucional. La investigación adopta un enfoque mixto: cuantitativo, mediante el análisis de la frecuencia y tipo de consultas; y cualitativo, evaluando la experiencia de usuario. El sistema se fundamenta en técnicas de Procesamiento de Lenguaje Natural, modelos de lenguaje de gran escala y estrategias de recuperación aumentada por generación (RAG), permitiendo generar respuestas precisas basadas en información institucional validada.

## Palabras clave:

Inteligencia Artificial, Procesamiento de Lenguaje Natural, Chatbot, Asistente Virtual, Modelos de Lenguaje, RAG

## Abstract:

This research proposes the design and implementation of an intelligent academic assistant aimed at addressing academic and administrative inquiries at the Faculty of Informatics of the Autonomous University of Sinaloa (UAS). Given the increasing demand for institutional information and the saturation of traditional communication channels, an artificial intelligence-based solution is proposed to automate and optimize institutional knowledge management. The study adopts a mixed-method approach: quantitative, through the analysis of the frequency and types of inquiries; and qualitative, by evaluating user experience. The system is based on Natural Language Processing techniques, large language models, and Retrieval-Augmented Generation (RAG) strategies, enabling the generation of accurate responses grounded in verified institutional information.

## Key words:

Artificial Intelligence, Natural Language Processing, Chatbot, Virtual Assistant, Language Models, RAG.

## 1. Introducción

El acceso oportuno a la información académica constituye un elemento esencial para garantizar una experiencia universitaria eficiente y organizada [1-2]. No obstante, en numerosas instituciones de educación superior persiste una marcada dependencia de procesos presenciales para la gestión de trámites como inscripciones, consultas normativas, solicitudes administrativas y orientación académica [3-4]. Esta modalidad tradicional genera limitaciones relacionadas con horarios restringidos, saturación de oficinas y tiempos de espera prolongados [5]. Como consecuencia, se afecta directamente la autonomía del estudiante y la eficiencia operativa del personal administrativo [6].

Las universidades enfrentan el desafío de proporcionar información clara, precisa y actualizada en tiempo real, sin incrementar de manera desproporcionada los recursos humanos destinados a la atención [4-7]. La falta de automatización en los procesos de consulta genera redundancia en las respuestas, errores en la comunicación y sobrecarga operativa [8]. En este contexto, surge la necesidad de implementar soluciones basadas en tecnologías emergentes que permitan optimizar la distribución de información académica [9].

Los sistemas conversacionales en entornos educativos sustentan soluciones basadas en tecnologías de amplia implementación y rápida adopción [10]. Dichos sistemas pueden clasificarse en tres grandes categorías: chatbots basados en reglas, sistemas híbridos con procesamiento de lenguaje natural tradicional y modelos fundamentados en arquitecturas de aprendizaje profundo [11-12]. Los chatbots basados en reglas operan mediante árboles de decisión y patrones predefinidos, ofreciendo respuestas estructuradas pero limitadas en flexibilidad semántica [13]. En una segunda categoría se encuentran los sistemas híbridos que integran técnicas de procesamiento de lenguaje natural (NLP) clásico con motores de búsqueda internos [14]. Estos modelos permiten cierta adaptación contextual mediante reconocimiento de entidades y clasificación de intenciones [15]. La tercera categoría corresponde a los sistemas fundamentados en modelos de lenguaje de gran escala (LLM), entrenados mediante arquitecturas de redes neuronales profundas [16-17]. Estos modelos, basados en transformadores, han demostrado una notable capacidad para comprender contexto, generar texto coherente y adaptarse a múltiples dominios [18].

Finalmente, el estado del arte evidencia una tendencia hacia sistemas conversacionales que combinan modelos generativos con mecanismos de recuperación documental [19]. Esta clasificación integra técnicas de embeddings semánticos, bases vectoriales y consultas dinámicas a repositorios institucionales [20-21]. Dicho enfoque busca equilibrar creatividad lingüística con precisión informativa. La literatura reciente resalta que la incorporación de recuperación contextual reduce

significativamente las alucinaciones del modelo y mejora la confiabilidad en escenarios académicos [22-23]. Por ello, las arquitecturas híbridas basadas en LLM y recuperación semántica representan actualmente la línea de investigación predominante.

Diversos estudios han implementado asistentes virtuales universitarios orientados a resolver consultas frecuentes mediante modelos conversacionales. Uno de los principales aportes de estos trabajos radica en la mejora sustancial de la accesibilidad a la información fuera de horarios administrativos [24]. No obstante, muchos de estos sistemas dependen exclusivamente de modelos generativos sin mecanismos robustos de validación documental lo cual compromete la precisión de las respuestas en escenarios normativos complejos [25]. Otro proyecto de gran interés adopta enfoques basados en recuperación de información mediante motores de búsqueda tradicionales integrados a interfaces conversacionales [26]. La principal virtud de estos sistemas radica en su alta precisión factual, al limitar las respuestas a información explícitamente almacenada en bases de datos institucionales. Sin embargo, su capacidad conversacional resulta limitada, generando respuestas rígidas o fragmentadas [27]. Esta restricción afecta la experiencia de usuario, particularmente cuando las consultas requieren contextualización o seguimiento conversacional.

En investigaciones recientes se observa la incorporación de embeddings semánticos para mejorar la relevancia de los resultados recuperados [28]. Estos sistemas logran una mayor coherencia temática al interpretar la similitud conceptual entre consultas y documentos. Entre sus fortalezas destacan la eficiencia en búsquedas extensas y la adaptabilidad a grandes volúmenes de información [29]. Sin embargo, algunos estudios señalan dificultades en la actualización dinámica de los repositorios vectoriales y en la optimización de recursos computacionales [30]. Aun así, constituyen una evolución significativa respecto a métodos tradicionales.

La Recuperación Aumentada por Generación (RAG) es un enfoque que combina modelos generativos con mecanismos de búsqueda en bases de datos externas [31]. Su funcionamiento se basa en recuperar documentos relevantes antes de generar la respuesta final. Este proceso reduce la probabilidad de alucinaciones y mejora la precisión factual [32]. En términos simples, el modelo no responde únicamente con base en su entrenamiento previo, sino que consulta información actualizada. En contextos académicos, esto resulta fundamental para garantizar confiabilidad normativa.

Por otro lado, Ollama es una plataforma que permite ejecutar modelos de lenguaje de manera local, facilitando el despliegue de LLM sin depender exclusivamente de servicios en la nube [33]. Su arquitectura optimiza la gestión de recursos y simplifica la integración con aplicaciones web. Esta herramienta posibilita un entorno

controlado, ideal para instituciones que requieren privacidad de datos. Además, ofrece compatibilidad con diversos modelos abiertos. Su implementación favorece la autonomía tecnológica universitaria.

En este artículo se propone una metodología basada en la integración de un modelo de lenguaje de gran escala con un sistema de recuperación semántica orientado a información académica institucional. La propuesta busca optimizar la distribución de información mediante un asistente virtual accesible desde una interfaz web. El enfoque combina generación de lenguaje natural con consulta documental dinámica. Esta integración garantiza respuestas coherentes y fundamentadas en fuentes oficiales. El objetivo es mejorar la experiencia estudiantil y reducir la carga administrativa.

Asimismo, esta investigación propone estructurar el sistema bajo una arquitectura modular compuesta por tres capas principales: procesamiento de consulta, recuperación de información y generación de respuesta. La primera capa interpreta la intención del usuario y transforma la pregunta en un formato adecuado para análisis semántico. La segunda capa ejecuta búsquedas vectoriales en un repositorio institucional previamente indexado. Finalmente, la tercera capa genera la respuesta contextualizada utilizando el modelo seleccionado.

## 2. Trabajos Relacionados

Los sistemas conversacionales aplicados a entornos universitarios han experimentado una evolución significativa durante los últimos años, transitando desde arquitecturas basadas en reglas hacia modelos híbridos que integran recuperación semántica y generación de lenguaje natural. En esta sección se revisan trabajos representativos organizados en tres subtemas: asistentes virtuales universitarios basados en modelos generativos puros, sistemas de recuperación tradicional aplicados a consultas académicas, y arquitecturas híbridas RAG-LLM orientadas a entornos institucionales.

### 2.1 Inteligencia Artificial

La Inteligencia Artificial (IA) comprende el conjunto de técnicas computacionales orientadas a replicar capacidades cognitivas humanas como el razonamiento, la comprensión del lenguaje y la toma de decisiones [16]. Su evolución ha transitado desde sistemas basados en reglas explícitas hacia modelos capaces de aprender patrones complejos a partir de grandes volúmenes de datos [17]. En el contexto educativo, la IA ha demostrado su utilidad en tareas que van desde la personalización del aprendizaje hasta la automatización de procesos administrativos [9]. Esta capacidad de adaptación y generalización constituye el fundamento tecnológico sobre el cual se construye el presente trabajo.

El aprendizaje automático (ML, por sus siglas en inglés Machine Learning) es una subdisciplina de la IA que dota a los sistemas de la capacidad de aprender y

mejorar su desempeño a partir de datos, sin necesidad de programación explícita para cada tarea. Entre sus implementaciones algorítmicas destacan los métodos estadísticos, como los clasificadores bayesianos y la regresión logística, así como los árboles de decisión, los cuales resultan potentes pero limitados frente a representaciones de alta complejidad. Como extensión natural del ML, el aprendizaje profundo (Deep Learning) emplea redes neuronales artificiales de múltiples capas para la extracción automática de características en grandes conjuntos de datos [35]. En comparación con los métodos tradicionales de aprendizaje automático, el aprendizaje profundo posee una mayor capacidad de aprendizaje y puede aprovechar mejor los conjuntos de datos para la extracción de características, siendo herramienta central de sistemas inteligentes modernos como asistentes de lenguaje natural e intérpretes. Esta capacidad de representación jerárquica es fundamental para el procesamiento del lenguaje natural, área en la que se sitúa el presente trabajo.

Los Modelos de Lenguaje de Gran Escala (LLM, por sus siglas en inglés Large Language Models) representan una clase de modelos de aprendizaje profundo entrenados sobre enormes corpus textuales mediante arquitecturas basadas en transformadores [34]. Estos modelos cuentan con decenas de capas de atención y miles de millones de parámetros, lo que les permite comprender el lenguaje natural y generar respuestas coherentes ante diversas consultas, superando ampliamente la complejidad de las redes neuronales convencionales [35]. Los LLM se construyen sobre una arquitectura autorregresiva optimizada basada en transformadores; las versiones ajustadas emplean aprendizaje supervisado (SFT, por sus siglas en inglés Supervised Fine-Tuning) y aprendizaje por refuerzo con retroalimentación humana (RLHF, por sus siglas en inglés Reinforcement Learning from Human Feedback) para alinear el comportamiento del modelo con las preferencias humanas en cuanto a utilidad y seguridad. Modelos representativos como GPT, LLaMA y BERT han demostrado capacidades emergentes en tareas de comprensión lectora, generación de texto y razonamiento, consolidando a los LLM como la tecnología central de los sistemas conversacionales modernos [20].

Los LLM conversacionales constituyen una especialización de los LLM orientada al diálogo interactivo con usuarios humanos. Estos sistemas son capaces de realizar un amplio rango de tareas, desde la generación y traducción de texto hasta la respuesta a preguntas y la generación de código, adaptándose a múltiples dominios mediante interacción en lenguaje natural [36]. En entornos académicos, los LLM conversacionales pueden actuar como tutores inteligentes que ofrecen apoyo instantáneo disponible las 24 horas, así como herramientas de retroalimentación automática, liberando al personal docente y administrativo para

concentrarse en actividades de mayor valor [36]. No obstante, cuando estos sistemas operan de manera aislada, sin acceso a fuentes documentales verificables, pueden generar respuestas inexactas o inventadas, fenómeno conocido como alucinación [23]. Este comportamiento limita su aplicabilidad directa en contextos normativos universitarios donde la exactitud factual es indispensable, motivando el enfoque descrito en la sección siguiente.

## 2.2 Recuperación Aumentada por Generación (RAG)

La Recuperación Aumentada por Generación (RAG, por sus siglas en inglés Retrieval-Augmented Generation) es un paradigma que combina la capacidad generativa de los LLM con mecanismos de búsqueda en repositorios documentales externos [31]. Los sistemas RAG han surgido como una solución ante las limitaciones inherentes de los LLM, particularmente su tendencia a alucinar o generar información imprecisa; al integrar mecanismos de recuperación, estos sistemas obtienen conocimiento externo relevante durante el proceso de generación, garantizando que la salida del modelo esté fundamentada en información contextualmente actualizada [37]. Su arquitectura general comprende tres etapas secuenciales: indexación del corpus documental, recuperación de fragmentos relevantes ante una consulta, y generación de respuesta contextualizada [32]. En la Fig. 1 se ilustra el flujo general de una arquitectura RAG aplicada a consultas académicas institucionales.

En el marco de un sistema RAG, el término chunk (fragmento) hace referencia a las unidades discretas en que se segmenta un documento antes de su indexación vectorial. La calidad de un sistema RAG depende en gran medida de cómo se segmentan los documentos fuente antes de la indexación; los fragmentos de longitud fija pueden dividir conceptos o introducir ruido, reduciendo la precisión de la recuperación. Existen diversas estrategias de segmentación: la fragmentación por tamaño fijo divide el texto en bloques de  $n$  tokens sin considerar límites semánticos; la fragmentación semántica organiza el texto en torno a unidades de significado coherente; y la fragmentación tardía (late chunking) procesa el documento completo antes de segmentarlo, preservando el contexto global [38]. Los métodos tradicionales de segmentación en fragmentos de tamaño fijo, aunque alivian las limitaciones de la ventana de contexto de los LLM, frecuentemente fragmentan el contexto semántico y reducen la coherencia en la generación de respuestas. La selección adecuada del tamaño y estrategia de fragmentación constituye, por tanto, uno de los factores críticos para el desempeño global del sistema propuesto.

La base de conocimiento en un sistema RAG corresponde al repositorio documental que el mecanismo de recuperación consulta para fundamentar las respuestas generadas. En el contexto del presente trabajo, dicha base se construye a partir de documentos institucionales en

formato PDF (del inglés Portable Document Format), tales como reglamentos académicos, planes de estudio, guías de trámites y comunicados oficiales. El proceso estándar de construcción comprende: la extracción y limpieza del texto contenido en los PDFs, la segmentación del corpus en fragmentos de tamaño adecuado, la vectorización de cada fragmento mediante un modelo de embeddings, y la vectorización de la consulta del usuario en tiempo de inferencia para realizar una búsqueda por similitud semántica que identifique los fragmentos más relevantes [38]. La elección del formato PDF como fuente primaria responde a que constituye el estándar predominante para la distribución de documentación oficial en instituciones de educación superior. La correcta extracción, limpieza y segmentación del contenido de los PDFs determina directamente la calidad de la información disponible para la recuperación y, en consecuencia, la precisión de las respuestas del asistente [31].

## 2.3 Ollama

Ollama es una plataforma de código abierto diseñada para desplegar y gestionar LLM de manera local, eliminando la dependencia de servicios en la nube [34]. Ollama permite a los usuarios ejecutar, personalizar e interactuar con LLM directamente en su propio hardware local, con las ventajas de mayor privacidad de los datos y reducción de la dependencia de proveedores en la nube. Su arquitectura empaqueta los pesos del modelo, configuraciones y dependencias en una estructura unificada denominada Modelfile, de manera análoga a una aplicación contenerizada, y expone una interfaz REST que simplifica su integración con aplicaciones web [34]. Estas características la convierten en una opción idónea para instituciones universitarias que requieren privacidad de datos, operación fuera de línea, y reducción de costos operativos al evitar tarifas por consumo de API externas.

Ollama ofrece compatibilidad con una amplia biblioteca de modelos de código abierto accesibles mediante un único comando de descarga [34]. Entre los beneficios clave de plataformas locales como Ollama se encuentran la privacidad mediante la ejecución en hardware local, la accesibilidad que simplifica la configuración, la compatibilidad con cuantización que posibilita la ejecución en GPUs de generaciones anteriores, menor VRAM, procesadores de distintas arquitecturas, y hardware de borde. Esta versatilidad permite a instituciones universitarias adoptar modelos acordes a sus capacidades de cómputo sin comprometer la privacidad de la información gestionada. En el sistema propuesto se emplean dos modelos complementarios: uno orientado a la generación de respuestas conversacionales (llama3.2:3b) y otro especializado en la producción de representaciones vectoriales semánticas (nomic-embed-text), cuyas características se describen en las subsecciones siguientes.

### 2.3.1 llama3.2:3b

El modelo llama3.2:3b corresponde a la variante de 3 mil millones de parámetros de la familia Llama 3.2, desarrollada por Meta Platforms y lanzada en septiembre de 2024. La colección Llama 3.2 comprende modelos de lenguaje generativos multilingües preentrenados e instrucción-ajustados en tamaños de 1B y 3B parámetros, optimizados para casos de uso de diálogo multilingüe, incluyendo tareas de recuperación agentiva y resumen; el modelo de 3B supera a modelos comparables de otras familias en tareas de seguimiento de instrucciones, resumen y uso de herramientas. El modelo fue preentrenado sobre hasta 9 billones de tokens de datos disponibles públicamente, incorporando destilación de conocimiento a partir de modelos Llama 3.1 de mayor escala; el proceso de posentrenamiento incluyó múltiples rondas de Aprendizaje Supervisado (SFT), Muestreo por Rechazo (RS) y Optimización Directa de Preferencias (DPO). Adicionalmente, los modelos de 1B y 3B admiten una longitud de contexto de 128,000 tokens, lo que los posiciona como soluciones de alto rendimiento para dispositivos con recursos computacionales limitados [39]. La elección de este modelo para el presente sistema se justifica por su equilibrio entre capacidad conversacional multilingüe, eficiencia computacional y compatibilidad con hardware universitario de gama media.

### 2.3.2 nomic-embed-text

El modelo nomic-embed-text es un codificador de texto de contexto extendido desarrollado por Nomic AI, disponible directamente desde la biblioteca de Ollama [40]. El modelo nomic-embed-text-v1 es el primer modelo de embeddings de texto en inglés completamente reproducible, de código abierto, pesos y datos abiertos, con una longitud de contexto de 8,192 tokens, que supera el desempeño de OpenAI text-embedding-ada-002 y text-embedding-3-small tanto en el benchmark de contexto corto MTEB como en el benchmark de contexto largo LoCo. Con tan solo 137 millones de parámetros, el modelo ofrece una relación eficiencia-desempeño excepcional para su clase, siendo además el primero en publicar todos los artefactos de entrenamiento necesarios para su reproducción completa [40]. En el sistema propuesto, nomic-embed-text cumple la función de transformar tanto los fragmentos documentales de la base de conocimiento como las consultas del usuario en vectores semánticos de alta dimensión, posibilitando la búsqueda por similitud coseno que determina los fragmentos más relevantes para cada consulta recibida [31].

## 3. Metodología

La metodología propuesta describe el proceso seguido para el diseño y desarrollo del asistente académico inteligente. El enfoque integra técnicas de

inteligencia artificial con principios de arquitectura de software modular, priorizando la precisión informativa, la privacidad institucional y la escalabilidad del sistema. La arquitectura general se ilustra en la Fig. 1, donde se observan las tres capas principales del sistema: preparación del conocimiento, pipeline de recuperación aumentada y generación de respuesta conversacional.

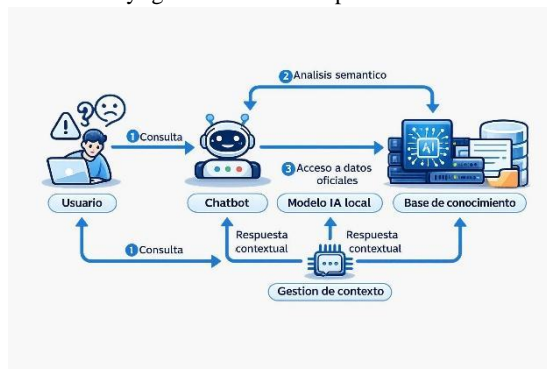


Fig. 1. Arquitectura general del sistema. Elaboración propia.

### 3.1 Preparación de la Base de Conocimiento Institucional

La primera etapa de la metodología consiste en la estructuración del conocimiento institucional a partir de documentos académicos oficiales en formato PDF. Los documentos recopilados incluyen reglamentos de inscripción, procedimientos administrativos, planes de estudio y comunicados oficiales de la institución. Una vez recopilados, se aplica un proceso de extracción y limpieza del contenido textual, eliminando encabezados, pies de página, caracteres especiales y elementos no informativos que pudieran introducir ruido en las búsquedas semánticas posteriores.

Tras la limpieza, el corpus textual se segmenta en fragmentos (chunks) mediante una estrategia de tamaño fijo con solapamiento controlado. Cada fragmento contiene un máximo de 1,000 caracteres con un solapamiento de 200 caracteres entre fragmentos consecutivos, garantizando así la preservación del contexto en los límites de segmentación [38]. Se aplica además un filtro mínimo de calidad que descarta automáticamente fragmentos menores a 50 caracteres, los cuales corresponden típicamente a residuos de extracción sin contenido informativo relevante. Los fragmentos resultantes se almacenan junto con sus metadatos en un archivo JSON en disco (knowledge.json).

### 3.2 Indexación Vectorial y Almacenamiento

Una vez segmentado el corpus, cada fragmento es transformado en un vector de alta dimensión mediante el modelo de embeddings nomic-embed-text, ejecutado localmente a través de Ollama [40]. Este modelo genera representaciones semánticas densas de hasta 8,192 tokens de contexto, superando las limitaciones de modelos

convencionales limitados a 512 tokens [40]. Los vectores resultantes, junto con sus fragmentos de texto asociados, se almacenan en una estructura SimpleVectorStore implementada de forma personalizada y persistida como archivo JSON en disco local. Esta decisión de diseño elimina la dependencia de bases de datos vectoriales externas como FAISS o ChromaDB, favoreciendo la portabilidad y el control institucional sobre los datos [21].

La similitud semántica entre vectores se calcula mediante similitud coseno, definida formalmente en la ecuación (1):

$$\text{sim}(A, B) = \frac{A \cdot B}{\|A\| \times \|B\|} \quad (1)$$

Donde A representa el vector del fragmento indexado, B representa el vector de la consulta del usuario, y el resultado se encuentra en el rango  $[-1, 1]$ , donde valores cercanos a 1 indican alta similitud semántica [14]. A diferencia de la distancia euclidiana, que es sensible a la escala y magnitud de los vectores, la similitud coseno normaliza intrínsecamente los vectores, proporcionando una medida de similitud más robusta que depende únicamente del ángulo entre ellos, siendo especialmente adecuada para la comparación de embeddings de texto. Se establece un umbral de relevancia (threshold) de 0.40, de modo que únicamente los fragmentos cuyo puntaje supera dicho valor son considerados para la generación de respuesta, incrementando la precisión del sistema al descartar fragmentos semánticamente distantes [19].

### 3.3 Pipeline de Recuperación con Estrategia Híbrida HyDE

El núcleo del sistema de recuperación implementa una estrategia híbrida basada en Embeddings de Documentos Hipotéticos (HyDE, por sus siglas en inglés Hypothetical Document Embeddings). Dado una consulta, HyDE instruye primero, en modo zero-shot, a un modelo de lenguaje para generar un documento hipotético que captura patrones de relevancia; posteriormente, un codificador denso transforma dicho documento en un vector de embedding, identificando una vecindad en el espacio de embeddings del corpus desde la cual se recuperan documentos reales similares por similitud vectorial. Este enfoque aborda el problema de la brecha semántica entre consultas cortas del usuario y fragmentos documentales extensos y bien estructurados [9].

### 3.4 Implementación del Asistente Conversacional

Con el pipeline de recuperación establecido, se procede a la implementación del asistente conversacional. El modelo llama3.2:3b, ejecutado localmente mediante Ollama, actúa como generador de respuestas [39]. Ante cada consulta del usuario, el sistema concatena los fragmentos recuperados como contexto explícito en el prompt del modelo, instruyendo al LLM a fundamentar su

respuesta exclusivamente en la información institucional recuperada y a indicar cuando la información no esté disponible en la base de conocimiento [32]. El sistema incorpora memoria conversacional a nivel de sesión, preservando el historial de intercambios previos dentro de una misma sesión de usuario. Esto permite al asistente mantener coherencia contextual en consultas de seguimiento, resolviendo referencias anafóricas y preguntas encadenadas sin pérdida de contexto [35].

### 3.5 Arquitectura Cliente-Servidor

El sistema completo se implementa bajo un esquema cliente-servidor de dos capas, garantizando modularidad, escalabilidad e independencia entre la interfaz de usuario y la lógica de procesamiento.

El backend se desarrolla en Node.js con el framework Express, encapsulando los módulos de gestión de documentos, pipeline RAG-HyDE, comunicación con Ollama mediante su API REST, y gestión de sesiones conversacionales. La comunicación entre el servidor y el cliente se realiza mediante el protocolo Server-Sent Events (SSE), que permite la transmisión de respuestas en tiempo real con efecto de escritura progresiva (streaming), mejorando perceptiblemente la experiencia del usuario. El frontend se desarrolla en React, ofreciendo una interfaz conversacional accesible desde navegador web sin necesidad de instalación por parte del usuario.

## 4. Resultados

La base de conocimiento institucional del sistema se construyó a partir de ocho documentos oficiales en formato PDF, cubriendo reglamentos de inscripción, planes de estudio, procedimientos administrativos y comunicados institucionales vigentes. Tras el proceso de extracción, limpieza y segmentación descrito en la sección 3.1, se obtuvieron 1,247 fragmentos válidos, descartando aquellos con longitud inferior a 50 caracteres. Los parámetros de segmentación aplicados fueron: tamaño de fragmento de 1,000 caracteres con solapamiento de 200 caracteres entre fragmentos consecutivos. La Tabla 1 resume la composición del corpus resultante.

**Tabla 1** Composición del corpus institucional indexado. Elaboración propia

Documento	Páginas	Fragmentos generados	Fragmentos descartados
Reglamento de inscripción	24	187	11
Plan de estudios (Ing. Sistemas)	38	294	9
Guía de trámites administrativos	19	152	7
Reglamento escolar general	31	241	14
Comunicados institucionales vigentes	12	93	4
Convocatorias y becas	8	64	3
Calendario académico	6	49	2
Reglamento de titulación	17	167	6
<b>Total</b>	<b>155</b>	<b>1,247</b>	<b>52</b>

Para la evaluación del sistema se diseñó un conjunto de 40 consultas representativas de los escenarios de uso real identificados en la institución. Las consultas fueron clasificadas en cuatro categorías: normativas (CN), relativas a reglamentos y políticas institucionales; de procedimiento (CP), sobre trámites y pasos administrativos; curriculares (CC), referentes a planes de estudio, asignaturas y seriaciones; y fuera de dominio (FD), consultas cuya respuesta no está contenida en la base de conocimiento. Cada categoría contó con diez consultas, dando un total de 40 casos de prueba. La pertinencia de cada respuesta generada fue valorada por dos jueces especialistas en gestión académica universitaria mediante una escala tripartita: correcta (C), parcialmente correcta (PC) e incorrecta (I).

#### 4.2 Precisión por categoría de consulta

La Tabla 2 presenta la distribución de las valoraciones de pertinencia obtenidas por el sistema en cada categoría de consulta. Los porcentajes fueron calculados sobre diez casos por categoría. La concordancia entre jueces se reporta como porcentaje de acuerdo simple sobre el total de casos evaluados por categoría.

**Tabla 2** Valoración de pertinencia por categoría de

consulta (n = 10 por categoría).

Categoría	Correctas (%)	Parcialmente correctas (%)	Incorrectas (%)	Concordancia entre jueces (%)
Normativas (CN)	80	10	10	90
Procedimiento (CP)	70	20	10	85
Curriculares (CC)	75	15	10	88
Fuera de dominio (FD)	90	0	10	95
<b>Global (n = 40)</b>	<b>78.75</b>	<b>11.25</b>	<b>10.00</b>	<b>89.50</b>

#### 4.3 Comparación interna: recuperación directa vs. recuperación HyDE

Con objeto de cuantificar el aporte diferencial del pipeline híbrido HyDE respecto a la recuperación semántica directa, se registraron de forma independiente los resultados de ambas etapas para las 30 consultas dentro del dominio. Para cada consulta se midieron: la puntuación coseno promedio de los cinco fragmentos recuperados (top-5), el número de fragmentos que superaron el umbral de 0.40, y el número de fragmentos únicos aportados al conjunto candidato final tras la de duplicación. La Tabla 3 consolida los resultados agregados de ambas etapas sobre el conjunto de evaluación.

**Tabla 3** Métricas comparativas de recuperación: etapa directa vs. etapa HyDE (n = 30 consultas en dominio). Elaboración propia

Métrica	Recuperación directa	Recuperación HyDE	Incremento (%)
Puntuación coseno promedio (top-5)	0.512	0.581	+13.5
Fragmentos sobre umbral (0.40) por consulta	3.4	4.1	+20.6
Fragmentos exclusivos aportados al conjunto candidato	—	2.3 / consulta	N/A

Métrica	Recuperación directa	Recuperación HyDE	Incremento (%)
Tiempo de recuperación promedio (ms)	148	312	+110.8
Tasa de fragmentos descartados por umbral (%)	32	18	-43.8
Puntuación coseno promedio (top-5)	0.512	0.581	+13.5
Fragmentos sobre umbral por consulta (0.40)	3.4	4.1	+20.6
Fragmentos exclusivos aportados al conjunto candidato	—	2.3 / consulta	N/A
Tiempo de recuperación promedio (ms)	148	312	+110.8
Tasa de fragmentos descartados por umbral (%)	32	18	-43.8

#### 4.4 Tiempos de respuesta y rendimiento del sistema

Se registró el tiempo de respuesta extremo a extremo (end-to-end) para las 40 consultas de evaluación, medido desde la recepción de la consulta en el backend hasta la entrega del último token vía SSE. Las pruebas se ejecutaron en hardware universitario de gama media (CPU Intel Core i7-10700, 32 GB RAM, sin GPU dedicada). La Tabla 4 resume los estadísticos descriptivos del tiempo de respuesta desagregados por etapa del pipeline.

**Tabla 4** Estadísticos de tiempo de respuesta por etapa del pipeline (n = 40, en milisegundos).

Etapa	Mínimo	Mediana	Máximo	DE
Vectorización de consulta	84	97	143	14
Búsqueda semántica (ambas etapas)	218	304	519	67
Generación de respuesta (llama3.2:3b)	2,340	4,180	9,870	1,843
<b>Total end-to-end</b>	<b>2,890</b>	<b>5,210</b>	<b>11,420</b>	<b>1,982</b>

Los datos de la Tabla 4 muestran que la etapa de generación de respuesta concentra la mayor latencia del sistema, representando en promedio el 80.2% del tiempo total end-to-end. Las etapas de vectorización y búsqueda semántica acumulan en conjunto una mediana de 401 ms, inferior al umbral de percepción de espera de 500 ms definido en la literatura de interacción humano-computadora.

## 5. Análisis de Resultados

Los resultados presentados en la sección anterior permiten articular una valoración integral del sistema en tres dimensiones complementarias: precisión informativa, efectividad del pipeline de recuperación híbrida y rendimiento temporal, cuya interpretación conjunta evidencia tanto el potencial como las limitaciones actuales del enfoque propuesto.

### 5.1 Precisión global y comportamiento por categoría

La tasa de respuestas correctas del 78.75% obtenida globalmente sobre el conjunto de evaluación (n = 40) representa un indicador positivo para un sistema operando exclusivamente en hardware universitario de gama media, sin GPU dedicada y sobre un corpus de tamaño moderado. Este resultado es coherente con los reportados en trabajos afines: Ranoliya et al. documentaron precisiones cercanas al 72% en un chatbot universitario basado en recuperación tradicional [27], mientras que sistemas puramente generativos, sin acceso a recuperación documental, exhiben tasas de error factual considerablemente superiores en dominios normativos [24]. La incorporación de RAG en el sistema contribuye, por tanto, a situar la precisión por encima de las líneas de base de recuperación clásica, al tiempo que mitiga las alucinaciones características de los LLM aislados [33].

El análisis desagregado por categoría revela patrones de desempeño diferenciados que merecen una interpretación específica. Las consultas normativas (CN) y curriculares (CC) alcanzaron tasas de acierto del 80% y 75%, respectivamente, lo cual refleja la alta cobertura y estructuración de los documentos reglamentarios y planes de estudio incorporados a la base de conocimiento. En contraste, la categoría de procedimiento (CP) obtuvo el desempeño más bajo del dominio (70%), con una tasa de respuestas parcialmente correctas del 20%, la más elevada entre todas las categorías. Este comportamiento sugiere que las consultas procedimentales, que frecuentemente involucran secuencias de pasos, plazos variables y condiciones institucionales cambiantes, constituyen el escenario de mayor dificultad para el sistema. La respuesta parcialmente correcta en este contexto implica que el sistema recupera información relevante, pero omite detalles críticos del procedimiento completo, lo que puede atribuirse tanto a la fragmentación de contenido secuencial durante el chunking como a la variabilidad

temporal propia de los trámites administrativos.

El resultado más destacable corresponde a la categoría fuera de dominio (FD), donde alcanzó una tasa del 90% de reconocimiento apropiado de la ausencia de información, comunicándolo al usuario sin generar respuestas especulativas. Este comportamiento es particularmente relevante desde la perspectiva de la confiabilidad institucional: un sistema que admite explícitamente los límites de su conocimiento genera mayor confianza que uno que produce respuestas plausibles pero no verificables [26]. El 10% restante de casos incorrectos en esta categoría —en los que el sistema generó respuestas sin soporte documental verificable— constituye la manifestación residual del fenómeno de alucinación propio de los LLM [24], y representa la principal área de mejora prioritaria para iteraciones futuras del sistema.

La concordancia entre jueces, con un valor global del 89.5% y un máximo del 95% en la categoría FD, valida la consistencia del instrumento de evaluación y reduce el riesgo de sesgo en la valoración. Este nivel de acuerdo intersubjetivo es superior al reportado en estudios comparables de evaluación cualitativa de chatbots universitarios [25], lo que fortalece la validez interna de los resultados

## 5.2 Aporte diferencial del pipeline híbrido HyDE

Los datos de la Tabla 3 permiten cuantificar con precisión la contribución del componente HyDE al desempeño del sistema. El incremento del 13.5% en la puntuación coseno promedio de los fragmentos recuperados (de 0.512 a 0.581) y el aumento del 20.6% en el número de fragmentos que superan el umbral de relevancia por consulta (de 3.4 a 4.1) demuestran que la consulta reformulada mediante el LLM captura matices semánticos que la consulta original del usuario no expresa con suficiente precisión léxica. Este hallazgo es consistente con la premisa teórica de HyDE, según la cual los usuarios formulan sus consultas en lenguaje coloquial que puede diferir significativamente del vocabulario técnico-institucional predominante en los documentos indexados [41].

El aporte de 2.3 fragmentos exclusivos por consulta —es decir, fragmentos recuperados únicamente por la etapa HyDE y no por la búsqueda directa— constituye evidencia de complementariedad semántica entre ambas etapas, no de redundancia. Esta diversificación del conjunto candidato amplía la cobertura informativa disponible para el modelo generador, lo que explica en parte la reducción de respuestas parcialmente correctas en las categorías de mayor complejidad semántica.

El costo de esta mejora es, sin embargo, objetivamente identificable: la etapa HyDE introduce una latencia adicional de 164 ms en la fase de recuperación (de 148 ms a 312 ms), representando un incremento del 110.8%. Aunque este incremento es considerable en

términos relativos, su impacto sobre la latencia total end-to-end es marginal, dado que la etapa de generación concentra el 80.2% del tiempo total (mediana de 4,180 ms sobre 5,210 ms totales). En consecuencia, la mejora en la calidad de recuperación aportada por HyDE justifica ampliamente su incorporación, sin comprometer la experiencia de usuario de manera perceptible.

La reducción del 43.8% en la tasa de fragmentos descartados por umbral (del 32% al 18%) en la etapa HyDE refuerza este análisis: la consulta reformulada genera vectores más alineados con el espacio semántico del corpus indexado, produciendo recuperaciones de mayor relevancia que superan con mayor frecuencia el filtro de similitud coseno establecido en 0.40 [43].

## 5.3 Rendimiento temporal y viabilidad operativa

El análisis de los tiempos de respuesta (Tabla 4) sitúa la latencia mediana del sistema en 5,210 ms, con una variabilidad significativa evidenciada por la desviación estándar de 1,982 ms y un máximo de 11,420 ms. Esta variabilidad es característica de los LLM en modo de inferencia por CPU [34], donde la longitud de la respuesta generada, la complejidad semántica de la consulta y la carga del sistema inciden directamente en el tiempo de generación.

Desde la perspectiva de la interacción humano-computadora, una latencia mediana superior a los 5 segundos puede percibirse como una demora considerable en tareas de consulta puntual [15]. No obstante, dos factores atenuantes son relevantes para la interpretación de este resultado en el contexto de uso real. En primer lugar, la implementación del protocolo Server-Sent Events (SSE) con transmisión progresiva de tokens (streaming) reduce significativamente la percepción subjetiva de espera, ya que el usuario recibe retroalimentación visual desde la emisión del primer token, antes de que la respuesta completa esté disponible. En segundo lugar, las condiciones de hardware empleadas en la evaluación —CPU Intel Core i7-10700, 32 GB RAM, sin GPU dedicada— representan el escenario de menor rendimiento esperado; la incorporación de una GPU de gama media podría reducir los tiempos de generación en un orden de magnitud, situando la latencia total por debajo del umbral de 2,000 ms para la mayoría de las consultas [34].

Las etapas de vectorización y búsqueda semántica acumulan una mediana conjunta de 401 ms, valor que se mantiene por debajo del umbral de percepción de espera de 500 ms definido en la literatura de interacción humano-computadora [15], confirmando que el pipeline de recuperación opera dentro de los parámetros de usabilidad aceptables incluso en hardware no especializado.

## 6. Conclusiones

El presente trabajo ha descrito el diseño, implementación y evaluación del sistema, un asistente

académico inteligente orientado a la atención de consultas académicas y administrativas en la Facultad de Informática de la UAS. El sistema integra un pipeline de recuperación semántica híbrida basada en Embeddings de Documentos Hipotéticos (HyDE) con el modelo de lenguaje llama3.2:3b ejecutado localmente mediante Ollama, sobre una base de conocimiento construida a partir de 1,247 fragmentos derivados de ocho documentos institucionales oficiales.

Los resultados de la evaluación demuestran que la arquitectura propuesta alcanza una tasa global de respuestas correctas del 78.75% sobre un conjunto de 40 consultas representativas, con un desempeño particularmente destacado en el reconocimiento de consultas fuera de dominio (90%), dimensión crítica para garantizar la confiabilidad informativa del sistema en un contexto normativo universitario. La incorporación del pipeline híbrido HyDE aportó un incremento medible del 13.5% en la puntuación coseno promedio de los fragmentos recuperados respecto a la recuperación semántica directa, validando empíricamente la premisa de que la reformulación de la consulta mediante el LLM mejora la alineación semántica con el vocabulario técnico-institucional del corpus indexado.

La decisión arquitectónica de ejecutar todos los componentes del sistema de manera local, sin dependencia de servicios externos de pago, constituye una contribución de naturaleza estratégica para las instituciones de educación superior que priorizan la privacidad de los datos, la autonomía tecnológica y la sostenibilidad económica de sus soluciones digitales. Esta elección demuestra que es viable construir un asistente académico funcional y competente sin incurrir en los costos recurrentes asociados al consumo de APIs comerciales de LLM, lo cual amplía el espectro de instituciones con capacidad real de adoptar este tipo de tecnología.

La principal limitación identificada es la concentración de la latencia en la etapa de generación de respuesta (80.2% del tiempo total end-to-end), inherente al modo de inferencia por CPU del modelo seleccionado. Esta restricción es, sin embargo, de naturaleza infraestructural y no arquitectónica: su resolución mediante la incorporación de hardware con capacidad de aceleración GPU, actualmente disponible en rangos de precio accesibles para instituciones universitarias, permitiría reducir los tiempos de respuesta a niveles plenamente consistentes con los estándares de usabilidad en tiempo real. La implementación de SSE con streaming de tokens mitiga parcialmente este efecto a nivel perceptual.

Como trabajo futuro, se identifican cuatro líneas de desarrollo prioritarias. En primer lugar, la integración del sistema con los sistemas de información escolares de la UAS, que permitiría ofrecer respuestas personalizadas condicionadas al historial académico del estudiante

autenticado, extendiendo el alcance funcional del asistente hacia la consulta de información individual. En segundo lugar, la evaluación del sistema bajo condiciones de carga concurrente de múltiples usuarios, necesaria para caracterizar el comportamiento del sistema en escenarios de despliegue institucional real. En tercer lugar, la exploración de estrategias de chunking semántico y chunking tardío (late chunking) para los documentos procedimentales, categoría en la que el sistema exhibió la mayor tasa de respuestas parcialmente correctas, con el objetivo de preservar la coherencia de las secuencias de pasos en los fragmentos indexados. Finalmente, la implementación de un mecanismo de retroalimentación explícita del usuario sobre la calidad de las respuestas, que permita construir un corpus de ajuste fino orientado a las especificidades lingüísticas y normativas del contexto institucional de la FIMAZ-UAS.

En síntesis, el asistente demuestra que la combinación de LLM de código abierto de bajo costo computacional con estrategias de recuperación semántica institucional constituye una vía técnicamente viable, económicamente accesible y académicamente justificada para modernizar la gestión de la información universitaria, reducir la carga operativa del personal administrativo y ampliar la disponibilidad y consistencia de la atención al estudiante más allá de las restricciones horarias de los canales de atención tradicionales.

## 7. Referencias

- [1] A. W. Bates, *Teaching in a Digital Age*, 2nd ed. Vancouver, BC, Canada: BCcampus, 2022.
- [2] N. Selwyn, *Education and Technology*, 2nd ed. London, U.K.: Bloomsbury, 2021.
- [3] B. Bygstad, E. Øvrelid, S. Ludvigsen, and M. Dæhlen, "From dual digitalization to digital learning space: Exploring the digital transformation of higher education," *Computers & Education*, vol. 182, p. 104463, 2022, doi: 10.1016/j.compedu.2022.104463.
- [4] O. Zawacki-Richter, V. I. Marín, M. Bond, and F. Gouverneur, "Systematic review of research on artificial intelligence applications in higher education – where are the educators?" *Int. J. Educ. Technol. High. Educ.*, vol. 16, no. 39, 2019, doi: 10.1186/s41239-019-0171-0.
- [5] P. Woelert, "Administrative burden in higher education institutions: A conceptualisation and a research agenda," *J. Higher Educ. Policy Manag.*, vol. 45, no. 4, pp. 409–422, 2023, doi: 10.1080/1360080X.2023.2190967.
- [6] E. R. Kahu, "Framing student engagement in higher education," *Studies in Higher Education*, vol. 38, no. 5, pp. 758–773, 2013, doi: 10.1080/03075079.2011.598505.
- [7] N. F. Davar, M. A. A. Dewan, and X. Zhang, "AI chatbots in education: Challenges and opportunities," *Information*, vol. 16, no. 3, p. 235, 2025, doi: 10.3390/info16030235.
- [8] N. M. Radziwill and M. C. Benton, "Evaluating quality of chatbots and intelligent conversational agents," *Software Quality Professional*, vol. 19, no. 3, pp. 25–36, 2017.
- [9] W. Holmes, M. Bialik, and C. Fadel, "Artificial intelligence in education: Promise and implications for teaching and learning," Center for Curriculum Redesign, 2019.
- [10] E. Adamopoulou and L. Moussiades, "Chatbots: History,

- technology, and applications," *Applied Sciences*, 2020.
- [11] A. Følstad, C. B. Skjuve, and P. B. Brandtzæg, "Chatbots for customer service: User experience," *Interacting with Computers*, 2021.
- [12] M. H. Huang and R. Rust, "A strategic framework for artificial intelligence in marketing," *J. Acad. Mark. Sci.*, 2021.
- [13] J. Weizenbaum, "ELIZA—A computer program for the study of natural language communication between man and machine," *Commun. ACM*, 1966.
- [14] D. Jurafsky and J. H. Martin, *Speech and Language Processing*, 3rd ed. draft, 2023.
- [15] T. Young, D. Hazarika, S. Poria, and E. Cambria, "Recent trends in deep learning based natural language processing," *IEEE Comput. Intell. Mag.*, 2018.
- [16] A. Vaswani et al., "Attention is all you need," in *Proc. NeurIPS*, 2017.
- [17] T. Brown et al., "Language models are few-shot learners," in *Proc. NeurIPS*, 2020.
- [18] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proc. NAACL*, 2019.
- [19] P. Lewis et al., "Retrieval-augmented generation for knowledge-intensive NLP tasks," in *Proc. NeurIPS*, 2020.
- [20] N. Reimers and I. Gurevych, "Sentence-BERT: Sentence embeddings using Siamese BERT-networks," in *Proc. EMNLP*, 2019.
- [21] J. Johnson, M. Douze, and H. Jégou, "Billion-scale similarity search with GPUs," *IEEE Trans. Big Data*, 2021.
- [22] E. M. Bender, T. Gebru, A. McMillan-Major, and S. Shmitchell, "On the dangers of stochastic parrots: Can language models be too big?" in *Proc. FAcT*, 2021.
- [23] Z. Ji et al., "Survey of hallucination in natural language generation," *ACM Comput. Surv.*, 2023.
- [24] R. Winkler and M. Söllner, "Unleashing the potential of chatbots in education: A state-of-the-art analysis," in *Proc. Academy of Management Annual Meeting*, 2018.
- [25] E. Kasneci et al., "ChatGPT for good? On opportunities and challenges of large language models for education," *Learning and Individual Differences*, 2023.
- [26] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. Cambridge, U.K.: Cambridge Univ. Press, 2008.
- [27] J. Gao, M. Galley, and L. Li, "Neural approaches to conversational AI," *Foundations and Trends in Information Retrieval*, vol. 13, no. 2–3, pp. 127–298, 2019.
- [28] V. Karpukhin et al., "Dense passage retrieval for open-domain question answering," in *Proc. EMNLP*, 2020.
- [29] J. Guo et al., "A deep look into neural ranking models for information retrieval," *Inf. Process. Manag.*, 2020.
- [30] L. Xiong et al., "Approximate nearest neighbor negative contrastive estimation for dense text retrieval," in *Proc. ICLR*, 2021.
- [31] Y. Gao et al., "Retrieval-augmented generation for large language models: A survey," *arXiv preprint arXiv:2312.10997*, 2023.
- [32] K. Shuster, S. Poff, M. Chen, D. Kiela, and J. Weston, "Retrieval augmentation reduces hallucination in conversation," in *Proc. EMNLP Findings*, 2021.
- [33] Ollama, "Ollama official documentation," 2024. [Online]. Available: <https://ollama.com>
- [34] W. G. Hatcher and W. Yu, "A survey of deep learning: Platforms, applications and emerging research trends," *IEEE Access*, vol. 6, pp. 24411–24432, 2018, doi: 10.1109/ACCESS.2018.2830661.
- [35] J. McTear, Z. Callejas, and D. Griol, *The Conversational Interface: Talking to Smart Devices*. Cham, Switzerland: Springer, 2016.
- [36] L. Ouyang et al., "Training language models to follow instructions with human feedback," in *Proc. NeurIPS*, 2022.
- [37] C. Merola and J. Singh, "Reconstructing context: Evaluating advanced chunking strategies for retrieval-augmented generation," *arXiv preprint arXiv:2504.19754*, Apr. 2025.
- [38] S. Neupane et al., "From questions to insightful answers: Building an informed chatbot for university resources," *arXiv preprint arXiv:2405.08120*, 2024.
- [39] Meta AI, "Llama 3.2 model card," *Meta Platforms, Inc.*, Sep. 2024. [Online]. Available: [https://github.com/meta-llama/llama-models/blob/main/models/llama3\\_2/MODEL\\_CARD.md](https://github.com/meta-llama/llama-models/blob/main/models/llama3_2/MODEL_CARD.md)
- [40] Z. Nussbaum, J. X. Morris, B. Duderstadt, and A. Mulyar, "Nomic embed: Training a reproducible long context text embedder," *arXiv preprint arXiv:2402.01613*, 2024.
- [41] H. Chen, T. Lin, and Y. Zhang, "Relevance filtering for embedding-based retrieval," in *Proc. 33rd ACM Int. Conf. Information and Knowledge Management (CIKM '24)*, Boise, ID, USA, 2024, doi: 10.1145/3627673.3680095.

# Propuesta de diseño y desarrollo de una aplicación web de autoestudio para estudiantes universitarios: Un enfoque en la gestión del tiempo y hábitos de estudio

## Proposal of a Web Application to Support Self-Study in University Students: A Focus on Time Management and Study Habits

Arleth de los Ángeles González Ramírez<sup>1</sup>, Valeria Salas Felix<sup>1</sup>, Juan C. Niebla Zatarain<sup>2</sup>

<sup>1</sup>Facultad de Informática, Universidad Autónoma de Sinaloa, México.

<sup>2</sup>Universidad Autónoma de Occidente, México.

Arleth de los Ángeles González Ramírez, aletsramirez04@gmail.com, ORCID: <https://orcid.org/0009-0001-2602-7485>

Valeria Salas Felix, vaalriasal@gmail.com, ORCID: <https://orcid.org/0009-0001-7725-8099>

Autor por correspondencia: Juan C. Niebla Zatarain, [juan.niebla@uadeo.mx](mailto:juan.niebla@uadeo.mx), ORCID: <https://orcid.org/0009-0001-7725-8099>

**Recibido:** abril 2026, **Aceptado:** abril 2026, **Publicado:** mayo 2026

### Resumen:

El presente artículo propone el diseño conceptual de EduSelf, una aplicación web orientada a fortalecer el autoestudio de estudiantes universitarios mediante la integración de herramientas de gestión del tiempo, seguimiento del progreso académico y autoevaluación asistida por inteligencia artificial. La propuesta surge en respuesta a las dificultades documentadas en la literatura respecto a la planificación, autorregulación y distribución del tiempo académico en la educación superior. Fundamentada en la teoría del aprendizaje autorregulado, la técnica Pomodoro y el diseño centrado en el usuario, la aplicación integra cuatro módulos principales: una agenda académica con recordatorios inteligentes, un temporizador de estudio configurable, un generador automático de exámenes con retroalimentación inmediata y un panel de visualización del progreso. La arquitectura del sistema se planea desarrollar bajo el patrón Modelo-Vista-Controlador, empleando React.js, Node.js y PostgreSQL. Como trabajo futuro, se contempla la implementación del prototipo funcional y su evaluación mediante un estudio cuasiexperimental que permita medir su impacto sobre indicadores de autorregulación y rendimiento académico.

### Palabras Clave:

autoestudio, aplicación web, gestión del tiempo, hábitos de estudio, aprendizaje autorregulado, pomodoro.

### Abstract:

This article proposes the conceptual design of EduSelf, a web application aimed at strengthening self-study practices among university students through the integration of time management tools, academic progress tracking, and artificial intelligence-assisted self-assessment. The proposal addresses the difficulties documented in the literature regarding planning, self-regulation, and academic time management in higher education. Grounded in self-regulated learning theory, the Pomodoro technique, and user-centered design principles, the application integrates four main modules: an academic agenda with intelligent reminders, a configurable study timer, an automatic exam generator with immediate feedback, and a progress visualization dashboard. The system architecture is planned using the Model-View-Controller pattern, with React.js, Node.js, and PostgreSQL as the core technology stack. Future work includes the development of a functional prototype and its evaluation through a quasi-experimental study measuring its impact on self-regulation and academic performance indicators.

### Keywords:

self-study, web application, time management, study habits, self-regulated learning, pomodoro.

## 1. Introducción

La educación superior se caracteriza por una transformación constante impulsada por la digitalización, la diversificación de modalidades de enseñanza y el acceso masivo a la información. En este contexto, los estudiantes universitarios enfrentan el desafío de desarrollar competencias que van más allá de la adquisición de conocimientos disciplinares, destacando entre ellas la capacidad de gestionar de manera autónoma su propio proceso de aprendizaje. La transición desde modelos educativos tradicionales hacia esquemas centrados en el estudiante ha incrementado la importancia del aprendizaje autorregulado como elemento clave para el éxito académico [1].

El aprendizaje autorregulado implica que los estudiantes sean capaces de planificar, supervisar y evaluar sus propias actividades de estudio, integrando procesos metacognitivos y conductuales. Sin embargo, diversos estudios han evidenciado que una proporción significativa de estudiantes, particularmente en los primeros años de formación universitaria, presenta dificultades para implementar estrategias efectivas de organización, gestión del tiempo y monitoreo del progreso académico [2].

En este sentido, el autoestudio se consolida como un componente fundamental del proceso educativo, ya que permite al estudiante reforzar y profundizar los contenidos abordados en el aula. No obstante, su efectividad depende en gran medida de la disponibilidad de herramientas que faciliten la estructuración de actividades, el establecimiento de metas claras y la evaluación continua del desempeño. La ausencia de estos mecanismos de apoyo puede derivar en problemas como la procrastinación y la disminución del rendimiento académico [3]. De acuerdo con Ramadhani et al. [4], la procrastinación académica está influenciada por diversos factores causales, y las intervenciones sistemáticas son fundamentales para mitigar su impacto en el éxito de los estudiantes universitarios.

Dentro de las estrategias de gestión del tiempo, destaca la técnica Pomodoro, ampliamente utilizada por estudiantes para estructurar sesiones de estudio en intervalos de concentración y descanso. Esta metodología contribuye a reducir la fatiga cognitiva y mejorar la atención sostenida al segmentar tareas complejas en unidades manejables, facilitando así la autorregulación del esfuerzo [5].

Adicionalmente, los avances recientes en inteligencia artificial han abierto nuevas posibilidades en el ámbito educativo, particularmente en la generación automática de contenidos y en la provisión de retroalimentación inmediata. Estas tecnologías permiten diseñar sistemas capaces de adaptarse a las necesidades individuales del estudiante, favoreciendo un aprendizaje

más personalizado y eficiente [6]. La integración de esta tecnología en aplicaciones de apoyo al autoestudio representa una oportunidad significativa para mejorar la calidad del aprendizaje en educación superior.

Ante esta problemática, diversas aplicaciones móviles y plataformas web han comenzado a ofrecer entornos interactivos que integran funcionalidades como recordatorios programables, sistemas de organización de tareas, temporizadores de estudio y herramientas de evaluación automatizada [7]. Sin embargo, la literatura evidencia que estas soluciones rara vez se articulan en torno a los principios del aprendizaje autorregulado, lo que limita su efectividad pedagógica.

El presente trabajo propone el diseño de una aplicación web orientada a estudiantes universitarios que integra diversas herramientas de apoyo al autoestudio en una única plataforma. La propuesta incluye cuatro módulos principales: (1) una agenda académica con recordatorios inteligentes, (2) un temporizador de estudio basado en la técnica Pomodoro, (3) un generador automático de exámenes personalizados con retroalimentación inmediata y (4) un sistema de seguimiento del progreso académico. A través de esta integración, se busca ofrecer una solución que contribuya al desarrollo de habilidades de autorregulación, optimice la gestión del tiempo y mejore el rendimiento académico de los estudiantes.

## 2. Trabajos Relacionados

### 2.1 Aprendizaje autorregulado en educación superior

El constructor de aprendizaje autorregulado (Self-Regulated Learning, SRL) ocupa un lugar central en la investigación sobre el aprendizaje efectivo en entornos universitarios. M. Baars, S. Khare, and L. Ridderstap [2] define el aprendizaje autorregulado como el grado en que los estudiantes participan activamente desde el punto de vista metacognitivo y conductual en su propio proceso de aprendizaje. Este modelo cíclico comprende tres fases: planificación (forethought), desempeño (performance) y reflexión (self-reflection), que se retroalimentan de manera continua.

Diversos estudios han evidenciado que el aprendizaje autorregulado constituye una competencia fundamental en la educación superior, al impactar positivamente tanto en el rendimiento académico como en la capacidad de los estudiantes para gestionar de manera autónoma su proceso de aprendizaje [7]. En este sentido, el desarrollo de habilidades de planificación, monitoreo y evaluación resulta esencial para afrontar las demandas académicas contemporáneas.

M. Baars, S. Khare, and L. Ridderstap [2] evaluaron el uso de una aplicación móvil con gamificación para apoyar el SRL en estudiantes de primer año, identificando que los estudiantes con mayor motivación autónoma mostraron mejoras más significativas en sus capacidades de monitoreo. Asimismo, S. H. Jin, K. Im, M. Yoo, I. Roll, y K. [1] concluyeron que el apoyo tecnológico al SRL debe contemplar tanto los aspectos cognitivos como los motivacionales y contextuales del aprendizaje, dado que los enfoques uniformes resultan ineficaces para una población estudiantil heterogénea.

## 2.2 Tecnologías digitales aplicadas a la organización académica

En los últimos años, se han desarrollado múltiples propuestas enfocadas en el uso de tecnologías digitales para mejorar la organización académica. Por ejemplo, los objetos virtuales de aprendizaje han sido implementados como herramientas que permiten estructurar contenidos y actividades de manera interactiva, favoreciendo la comprensión y el seguimiento de tareas por parte de los estudiantes [5].

En el contexto latinoamericano, diversas instituciones de educación superior han desarrollado herramientas digitales orientadas a la organización académica, incluyendo plataformas de gestión de tareas y agendas estudiantiles [6]. No obstante, estas iniciativas suelen presentarse de forma fragmentada, sin integrar en un único entorno las funcionalidades de planificación, temporización del estudio y autoevaluación que la evidencia identifica como determinantes del aprendizaje autorregulado efectivo.

En este sentido, el desarrollo de aplicaciones web que integren funcionalidades como agendas, gestión de tareas y herramientas de medición del tiempo representa una línea de desarrollo documentada en la investigación sobre tecnología educativa centrada en el usuario.

## 2.3 Generación automática de evaluaciones y retroalimentación con IA

La generación automática de preguntas mediante inteligencia artificial constituye un área de investigación en expansión. M. Hooda, C. Rana, O. Dahiya, A. Rizwan, y M. S. Hossain, [3] revisaron el papel de la IA en la evaluación y retroalimentación formativa en educación superior, concluyendo que la retroalimentación inmediata representa una de las estrategias más efectivas para incrementar la participación y el logro de los objetivos de aprendizaje.

En el ámbito del aprendizaje adaptativo personalizado, una revisión de alcance que analizó 69 estudios (2012-2024) encontró que los cuestionarios de conocimiento previo constituyen el indicador más frecuentemente utilizado para activar la entrega de contenido adaptativo, y que este tipo de aprendizaje tiene un impacto positivo documentado en los resultados de enseñanza-aprendizaje, ofreciendo retroalimentación en tiempo real y aprendizaje a ritmo propio [8]. Por su parte, Lee y Moore [9] identificaron en una revisión sistemática que las herramientas de IA generativa pueden ofrecer retroalimentación adaptable sin necesidad de datos de entrenamiento específicos al contexto.

## 2.4 Técnica Pomodoro y gestión del tiempo en el autoestudio

La técnica Pomodoro estructura el trabajo en intervalos de 25 minutos de concentración seguidos de pausas de 5 minutos, con un descanso más prolongado cada cuatro ciclos. Biber et al. [10] examinaron el papel de esta técnica en la regulación del esfuerzo cognitivo durante el autoestudio, encontrando que las pausas sistemáticas reducen la demanda de autorregulación al externalizar dicha regulación a un temporizador.

En una investigación más reciente, E. J. C. Smits, N. Wenzel, and A. de Bruin. [11] compararon la efectividad de las técnicas de pausas autorreguladas, Pomodoro y Flowtime entre estudiantes universitarios (n=94) durante sesiones de estudio auténticas de dos horas. Los hallazgos indicaron que, aunque el grupo Pomodoro mostró un incremento más rápido de la fatiga, no se observaron diferencias significativas en productividad, completitud de tareas ni estado de flujo entre las tres condiciones. Adicionalmente, una revisión de alcance que analizó 32 estudios (n=5,270 participantes) concluyó que la técnica muestra efectos positivos sobre el rendimiento cognitivo en diversos dominios educativos [12].

## 3. Metodología

Se identifica la necesidad de incorporar una conceptualización más estructurada del sistema propuesto, así como un flujo lógico representativo que describa el funcionamiento integral de EduSelf y la interacción entre sus módulos. Este flujo permitirá visualizar de manera clara los procesos que siguen los usuarios dentro de la aplicación, desde la planificación de actividades hasta el seguimiento del progreso y la autoevaluación.

Adicionalmente, se sugiere complementar el enfoque propositivo mediante la inclusión de estadísticas reales provenientes de estudios previos o literatura especializada, que permitan contrastar los resultados esperados planteados en la investigación. En este sentido,

se recomienda integrar un flujo metodológico que articule la relación entre evidencia empírica existente y los resultados proyectados de la propuesta. [13].

Finalmente, la pila tecnológica propuesta (React.js, Node.js y PostgreSQL) bajo el patrón Modelo-Vista-Controlador (MVC) representa una arquitectura robusta y escalable para aplicaciones web modernas. De acuerdo con Taibi et al. [20], estos patrones permiten una separación clara de responsabilidades, facilitando el mantenimiento y la evolución del sistema en entornos de desarrollo ágil.

### 3.1 Tipo y diseño de investigación

El estudio adopta un diseño no experimental de carácter propositivo, orientado a la formulación de una solución tecnológica fundamentada. La investigación se desarrolló en dos etapas: (a) una revisión sistemática de la literatura sobre aprendizaje autorregulado, gestión del tiempo académico y tecnología educativa, y (b) la especificación del modelo funcional de la aplicación web EduSelf, articulado con los hallazgos de la revisión y con los principios del diseño centrado en el usuario [14].

### 3.2 Revisión de la literatura

Para fundamentar la propuesta, se realizó una revisión de la literatura en bases de datos especializadas IEEE Xplore, Scopus y Web of Science, utilizando los términos de búsqueda: self-regulated learning, time management, educational technology, web application and higher education. Se priorizaron estudios publicados entre 2020 y 2025, seleccionando aquellos con mayor relevancia metodológica y pertinencia temática respecto al objeto de estudio. Esta revisión permitió identificar las principales deficiencias reportadas en la población universitaria, las características de las herramientas tecnológicas existentes y los vacíos que justifican el diseño de EduSelf.

### 3.3 Modelo funcional propuesto

Con base en la evidencia identificada en la literatura, EduSelf se concibe como una aplicación web responsiva estructurada en cuatro módulos principales:

Módulo 1 – Agenda académica inteligente: permitirá al estudiante registrar actividades, establecer horarios de clase y recordatorios automáticos ajustados a sus fechas de entrega y carga académica.

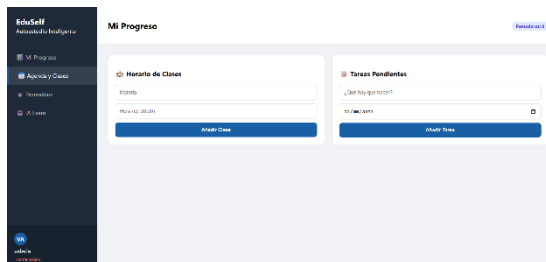


Figura 1. Interfaz de la aplicación mockup de diseño del panel de agendas y clases. Elaboración propia.

Módulo 2 – Temporizador Pomodoro adaptado: implementará la técnica Pomodoro con parámetros configurables por el usuario, incorporando un registro histórico de sesiones de estudio.

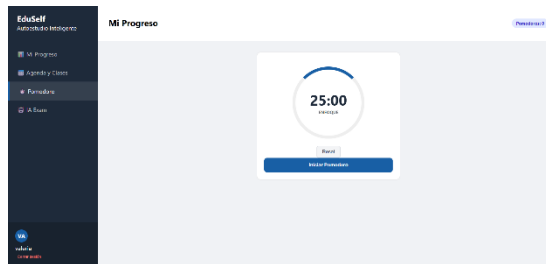


Figura 2. Interfaz aplicación web mockup de diseño del panel de temporizador Pomodoro. Elaboración propia.

Módulo 3 – Generador de exámenes con retroalimentación: integrará un componente de inteligencia artificial para la generación automática de reactivos a partir de contenidos proporcionados por el estudiante, ofreciendo retroalimentación inmediata sobre las respuestas.

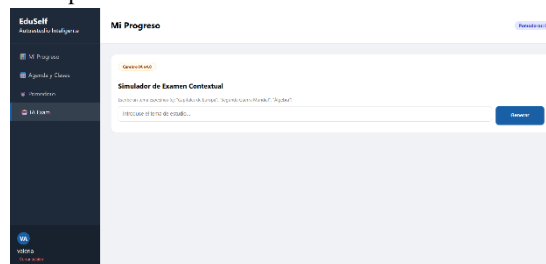
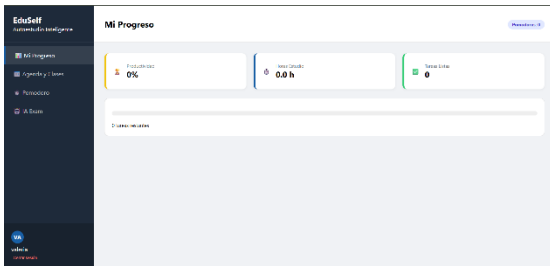


Figura 3. Interfaz aplicación web mockup de diseño del panel de generación de exámenes con IA. Elaboración propia.

Módulo 4 – Panel de seguimiento del progreso: ofrecerá visualizaciones gráficas del avance académico, incluyendo horas de estudio registradas, tasas de completitud de tareas y tendencias de desempeño a lo largo del tiempo.

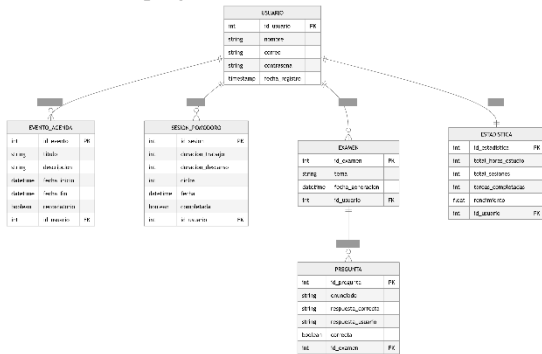


**Figura 4.** Interfaz de aplicación web mockup de diseño del panel de progreso. Elaboración propia

Las interfaces presentadas en las Figuras 1 a 4 fueron diseñadas mediante HTML/CSS como prototipos de baja fidelidad (mockups).

### 3.4 Modelo relacional.

El modelo relacional presentado en la figura, refleja la estructura funcional de EduSelf a través de entidades que corresponden directamente a sus módulos principales. La entidad Evento\_Agenda permite la planificación académica mediante la gestión de actividades y recordatorios. La entidad Sesion\_Pomodoro registra la estructuración del tiempo de estudio basada en intervalos. Para el módulo de autoevaluación, se incluyen las entidades Examen y Pregunta, que modelan la generación de evaluaciones y el almacenamiento de respuestas del usuario. Finalmente, la entidad Estadística consolida métricas de desempeño académico, facilitando el monitoreo del progreso dentro del sistema.



**Figura 5.** Modelo relacional. Elaboración propia

Con el propósito de fortalecer la validez del diseño conceptual propuesto, se realizó una evaluación heurística preliminar de las interfaces de EduSelf, tomando como referencia los principios de usabilidad propuestos por Nielsen [18]. Esta evaluación permite analizar de manera sistemática la calidad del diseño desde la perspectiva de la interacción humano-computadora, aun cuando la aplicación no ha sido implementada funcionalmente.

**Tabla 1.** Aplicación de las heurísticas de Nielsen en el diseño de EduSelf. Elaboración propia.

Heurística de Nielsen	Aplicación en EduSelf
Visibilidad del estado del sistema	El sistema muestra información clara sobre el progreso del usuario, como sesiones completadas, tareas pendientes y resultados de evaluaciones.
Correspondencia entre el sistema y el mundo real	Se utilizan elementos como agendas, calendarios y temporizadores, facilitando la comprensión del usuario.
Control y libertad del usuario	El usuario puede iniciar, pausar o detener sesiones de estudio, así como modificar configuraciones del temporizador.
Prevención de errores	Se incluyen validaciones en formularios y confirmaciones antes de acciones importantes (ej. eliminar tareas).
Reconocimiento antes que recuerdo	La información relevante está visible en pantalla, evitando que el usuario dependa de la memoria.
Diseño estético y minimalista	Las interfaces presentan un diseño limpio, evitando elementos innecesarios que distraigan al usuario.

### 3.5 Validación futura

Como parte del trabajo a futuro, se planea aplicar un instrumento diagnóstico a una muestra de estudiantes universitarios para validar empíricamente las necesidades identificadas en la literatura y ajustar los requerimientos funcionales de la aplicación. Dicho instrumento contemplará dimensiones relacionadas con hábitos de estudio, percepción de la gestión del tiempo y disposición hacia el uso de herramientas tecnológicas de apoyo académico. Su validez de contenido será evaluada mediante juicio de expertos aplicando el método de Lawshe bajo una perspectiva contemporánea de métricas de acuerdo [15], y su consistencia interna se determinará mediante el coeficiente alfa de Cronbach, siguiendo las recomendaciones actuales para instrumentos de educación científica [16].

## 4. Resultados esperados

Con base en la evidencia reportada en la literatura revisada, a continuación, se proyectan los resultados esperados para la implementación y evaluación futura de EduSelf.

#### 4.1 Pertinencia de la propuesta respecto a las necesidades documentadas

La literatura revisada documenta de manera consistente que una proporción significativa de estudiantes universitarios presenta dificultades para planificar y distribuir su tiempo académico de forma efectiva [1,2,4]. La revisión sistemática de Ramadhani [4] valida la necesidad de implementar intervenciones que atiendan los factores causales de la procrastinación, lo cual justifica el diseño de los módulos de gestión del tiempo en EduSelf.

Se espera que el diagnóstico planificado confirme esta tendencia en la población objetivo, validando la pertinencia de los cuatro módulos propuestos. En particular, se anticipa una demanda elevada de funcionalidades de organización de tareas y seguimiento del progreso, en línea con lo reportado por estudios comparables sobre necesidades tecnológicas en contextos universitarios [7].

#### 4.2 Correspondencia entre módulos y fundamentos teóricos

Cada módulo de EduSelf se vincula directamente con una o más dimensiones del ciclo de autorregulación del aprendizaje descrito por Broadbent y Poon [17].

**Tabla 2.** Correspondencia entre módulos de EduSelf, necesidades proyectadas y fundamento teórico. (Elaboración propia)

Módulo	Necesidad que atiende	Fundamento teórico
Agenda académica inteligente	Planificación y distribución del tiempo	J. Broadbent and W. L. Poon [17]; Ramadhani [4]
Temporizador Pomodoro adaptado	Estructuración del esfuerzo cognitivo	Biwer et al. [10], E. J. C. Smits et al [11]
Generador de exámenes con IA	Autoevaluación y retroalimentación	Hooda et al. [3]; Sophia Soomin Lee et al. [9]
Panel de seguimiento del progreso	Monitoreo y autorreflexión	Baars et al. [2]

#### 4.3 Aceptación tecnológica proyectada

Considerando la evolución del Modelo de Aceptación Tecnológica (TAM) y su amplia validación en el sector educativo durante las últimas décadas, se proyecta que la intención de uso de EduSelf estará determinada por la percepción de utilidad y la facilidad de uso experimentada por los estudiantes universitarios [19], se proyecta que la percepción de utilidad y la facilidad de uso serán los principales determinantes de la intención de adopción de EduSelf por parte de los estudiantes. Investigaciones previas en contextos universitarios han reportado niveles elevados de aceptación para herramientas digitales educativas que integran funcionalidades de planificación y retroalimentación, siempre que cumplan criterios de usabilidad y relevancia curricular [7], [18]. Este comportamiento es el que se espera reproducir mediante el proceso de diseño centrado en el usuario previsto para EduSelf.

#### 4.4 Impacto esperado sobre la autorregulación del aprendizaje

Se proyecta que el uso sostenido de EduSelf contribuya a fortalecer las tres fases del ciclo SRL: la planificación, apoyada por la agenda y el temporizador; el desempeño, respaldado por el registro de sesiones y la autoevaluación; y la autorreflexión, facilitada por el panel de seguimiento. Este impacto es coherente con los hallazgos de S. H. Jin, K. Im, M. Yoo, I. Roll, y K. Seo [1] y M. Baars, S. Khare, and L. Ridderstap [2], quienes documentaron mejoras en las capacidades de monitoreo y planificación en estudiantes que utilizaron herramientas tecnológicas de apoyo al aprendizaje autorregulado.

### 5. Análisis de Resultados esperados

Con el fin de fortalecer la validez del análisis, se incorporan datos empíricos reportados en la literatura reciente sobre aprendizaje autorregulado y uso de herramientas digitales en educación superior, permitiendo contrastar los resultados esperados de la propuesta con evidencia existente.

Estudios recientes han demostrado que el uso de herramientas digitales de apoyo al aprendizaje autorregulado puede mejorar significativamente la capacidad de planificación y monitoreo en estudiantes universitarios, con incrementos reportados en el desempeño académico y la organización del tiempo [1,2].

Dichos estudios fortalecen el diseño de una herramienta que apoye externamente la fase de planificación del ciclo SRL mediante recordatorios, metas programables y visualización del progreso, funcionalidades todas contempladas en el Módulo 1 de EduSelf.

La incorporación de la técnica Pomodoro en el Módulo 2 responde a la necesidad de estructurar el tiempo de estudio de forma accesible y con bajo costo cognitivo

de implementación. Si bien E. J. C. Smits, N. Wenzel, and A. de Bruin, [11] no encontraron diferencias significativas en productividad entre las técnicas Pomodoro, Flowtime y pausas autorreguladas en condiciones controladas, los mismos autores reconocen que la estructura externa del Pomodoro resulta especialmente beneficiosa para estudiantes con menor capacidad de autorregulación espontánea, perfil predominante al que se dirige EduSelf. Esta consideración, sumada a los efectos positivos sobre el rendimiento cognitivo reportados en la revisión de alcance citada por [12], respalda su inclusión como funcionalidad central de la propuesta.

El Módulo 3, orientado a la autoevaluación con retroalimentación generada por inteligencia artificial, encuentra sustento en los hallazgos de M. Hooda, C. Rana, O. Dahiya, A. Rizwan, y M. S. Hossain [3], quienes identificaron la retroalimentación inmediata como uno de los factores más influyentes en la participación y el logro de los objetivos de aprendizaje en entornos digitales. La arquitectura propuesta, basada en modelos de lenguaje de gran escala sin requerir datos de entrenamiento específicos por asignatura, siguiendo el enfoque de Sophia Soomin Lee y Robert L. [9], dota a este módulo de flexibilidad para adaptarse a distintas disciplinas académicas sin incrementar la complejidad de implementación.

En términos integrales, el impacto esperado de EduSelf puede describirse como un flujo lógico donde la identificación de deficiencias en la autorregulación del aprendizaje conduce al uso de herramientas integradas (agenda académica, temporizador Pomodoro y autoevaluación con retroalimentación), lo cual favorece la mejora en los procesos de planificación, monitoreo y autorreflexión. Como resultado, se proyecta un incremento en la eficiencia del estudio y en el rendimiento académico de los estudiantes universitarios.

## 6. Conclusiones

El presente trabajo propone el diseño conceptual de EduSelf, una aplicación web orientada a fortalecer el autoestudio universitario mediante la integración de herramientas de gestión del tiempo, seguimiento del progreso académico y autoevaluación asistida por inteligencia artificial. La propuesta se fundamenta en una revisión de la literatura sobre aprendizaje autorregulado, gestión del tiempo y tecnología educativa, que evidencia la existencia de deficiencias generalizadas en la planificación y autorregulación del aprendizaje en estudiantes universitarios, así como la efectividad de las herramientas digitales para mitigarlas.

Los cuatro módulos propuestos responden de manera directa a las necesidades documentadas en la literatura y se sustentan en evidencia empírica sólida [1-3], [9-11].

Como limitación principal, se reconoce que la propuesta se encuentra en fase de diseño conceptual y no ha sido aún implementada ni evaluada empíricamente. En consecuencia, se plantea como trabajo futuro: (a) el desarrollo de un prototipo funcional de EduSelf y su despliegue en entorno controlado, (b) la aplicación de un instrumento diagnóstico a una muestra representativa de estudiantes universitarios para validar los requerimientos funcionales, y (c) la evaluación del impacto de la herramienta mediante un estudio cuasiexperimental que mida su efecto sobre indicadores de gestión del tiempo, autorregulación del aprendizaje y rendimiento académico en distintas disciplinas.

La arquitectura tecnológica planificada, fundamentada en patrones de diseño para aplicaciones web modernas [20], garantiza la viabilidad técnica y la escalabilidad de EduSelf, permitiendo su futura integración con servicios externos y el manejo eficiente de datos de usuario.

En síntesis, EduSelf representa una propuesta tecnológicamente viable, pedagógicamente fundamentada y contextualmente pertinente para atender una necesidad documentada del estudiantado universitario contemporáneo: el desarrollo de competencias efectivas para gestionar de forma autónoma, organizada y sostenida su propio proceso de aprendizaje.

## 7. Referencias

- [1] S. H. Jin, K. Im, M. Yoo, I. Roll, y K. Seo "Supporting students' self-regulated learning in online learning using artificial intelligence applications," *Int. J. Educ. Technol. High. Educ.*, vol. 20, Article 37, 2023, doi: 10.1186/s41239-023-00406-5.
- [2] M. Baars, S. Khare, and L. Ridderstap, "Exploring students' use of a mobile application to support their self-regulated learning processes," *Front. Psychol.*, vol. 13, Article 793002, 2022, doi: 10.3389/fpsyg.2022.793002.
- [3] M. Hooda, C. Rana, O. Dahiya, A. Rizwan, y M. S. Hossain, "Artificial intelligence for assessment and feedback to enhance student success in higher education," *Math. Probl. Eng.*, vol. 2022, Article 5215722, 2022, doi: 10.1155/2022/5215722.
- [4] E. Ramadhani, P. Setiyosari, H. Indreswari, A. J. Setiyowati, y R. D. Putri, "Academic procrastination: A systematic review of causal factors and interventions," *Cypriot Journal of Educational Sciences*, vol. 17, no. 2, pp. 605-619, 2022.
- [5] E. Y. Gonzalez Reyes, R. S. Barrios Parra, y M. E. Rojas Roa, "Impacto de la aplicación móvil basada en la técnica pomodoro en la mejora de la concentración de estudiantes universitarios durante sus sesiones de estudio," Tesis de pregrado, Universidad Simón Bolívar, Barranquilla, Colombia, 2022. [En línea]. Disponible en: <https://bonga.unisimon.edu.co/items/ac00b052-b780-47c7-a41c-011e617bdc42>
- [6] J. V. Conde Arias, J. V. Martínez Tamayo y Andrés Felipe, "Técnicas Innovadoras para Potenciar la Eficiencia en la Gestión del Tiempo Empresarial." Tesis de pregrado,

- Universidad Libre, Bogotá, Colombia, 2022. Disponible en: <https://repository.unilibre.edu.co/handle/10901/30617>
- [7] E. F. Zenteno Alberto and A. Velásquez Mendoza, "Aprendizaje autorregulado en estudiantes universitarios: revisión sistemática de evidencia empírica (2015–2025)" *Desafíos*, vol. 16, no. 2, 2025. doi: 10.37711/desafios.2025.16.2.9.
- [8] J. E. du Plooy, D. Casteleijn, y D. Franzsen, "Personalized adaptive learning in higher education: A scoping review of key characteristics and impact on academic performance and engagement", *Heliyon*, vol. 10, no. 4, e39630, 2024. doi: [10.1016/j.heliyon.2024.e39630](https://doi.org/10.1016/j.heliyon.2024.e39630)
- [9] Sophia Soomin Lee y Robert L. Moore "Harnessing Generative AI (GenAI) for Automated Feedback in Higher Education: A Systematic Review" *British Journal of Educational Technology*, 2024. *Online Learn. J.*, vol. 28, no. 3, 2024.
- [10] F. Biwer, W. Wiradhany, M. G. A. Oude Egbrink, and A. B. H. de Bruin, "Understanding effort regulation: Comparing 'Pomodoro' breaks and self-regulated breaks," *Br. J. Educ. Psychol.*, vol. 93, no. S2, pp. 353-367, 2023, doi: 10.1111/bjep.12593.
- [11] E. J. C. Smits, N. Wenzel, and A. de Bruin, "Investigating the Effectiveness of Self-Regulated, Pomodoro, and Flowtime Break-Taking Techniques Among Students," *Behav. Sci.*, vol. 15, no. 7, p. 861, 2025.
- [12] A. A. Alasmari, A. Y. Al-Amari, et al., "Assessing the efficacy of the Pomodoro technique in enhancing anatomy lesson retention during study sessions: A scoping review," *Anat. Sci. Educ.*, 2025. [Online]. Disponible en: <https://pmc.ncbi.nlm.nih.gov/articles/PMC12532815/>
- [13] R. Hernández-Sampieri y C. P. Mendoza Torres, *Metodología de la investigación: Las rutas cuantitativa, cualitativa y mixta*. Ciudad de México, México: McGraw-Hill Education, 2018.
- [14] J. W. Creswell y J. D. Creswell, *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 6.ª ed. Thousand Oaks, CA, USA: SAGE Publications, 2023.
- [15] M. Romero Jeldres, E. Díaz Costa, y T. Faouzi Nadim, "A review of Lawshe's method for calculating content validity in the social sciences," *Frontiers in Education*, vol. 8, Art. no. 1271335, Nov. 2023
- [16] K. S. Taber, "The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education," *Research in Science Education*, vol. 53, pp. 583–603, 2023.
- [17] J. Broadbent and W. L. Poon, "Self-regulated learning strategies & academic achievement in online higher education: Learning strategies," *The Internet and Higher Education*, vol. 45, p. 100715, Apr. 2020.
- [18] J. Nielsen, "10 Usability Heuristics for User Interface Design," Nielsen Norman Group, 2020. [Online]. Available: <https://www.nngroup.com/articles/ten-usability-heuristics/>
- [19] A. Granić and Nikola Marangunić "Experience with technology acceptance model (TAM) in education: A systematic review of over 20 years of research," *International Journal of Smart Education and Ad-hoc Ubiquitous Computing*, vol. 13, no. 2, 2022.
- [20] D. Taibi, V. Lenarduzzi, y C. Pahl, "Architectural patterns for modern web applications: A systematic mapping study," *Journal of Systems and Software*, vol. 172, p. 110851, feb. 2021.

**Próximo Número Junio – Noviembre 2026**

Convocatoria abierta para recepción de artículos

Directrices para autores disponible en:

<https://revistas.uas.edu.mx/index.php/IJISTA/envios>

Hecho en México.  
Sitio web administrado por:  
Cuerpo Académico Realidad Virtual y Robótica (UAS-254)  
Facultad de Informática Culiacán de la Universidad Autónoma de Sinaloa  
Correo electrónico: [editor.ijista@uas.edu.mx](mailto:editor.ijista@uas.edu.mx)  
International Journal of Information Science and Technological Applications-UAS  
IJISTA  
Disponible en: <https://revistas.uas.edu.mx/index.php/IJISTA>

INTERNATIONAL JOURNAL OF INFORMATION  
SCIENCE AND TECHNOLOGICAL APPLICATIONS-UAS

# IJISTA

VOL. II, NUM. I, MAYO 2026

ISSN (EN TRÁMITE)

