LÍNEA IMAGINARIA
REVISTA DE INVESTIGACIÓN EN ESTUDIOS SOCIALES Y DE FRONTERA

Número 25. Vol. 1 (2026), pp. 25- 69/ Edición Continua

ISSN en Línea: 2477-9415 / ISSN Impreso: 2477-9415
Depósito Legal: pp 200602TA2254

REPORTE DE INVESTIGACIÓN

EVALUACIÓN DE CHATGPT Y GEMINI PARA APRENDIZAJE DE DESARROLLO WEB BÁSICO

Yhonny Mancilla Díaz
Orcid: 0009-0005-3302-9875
yhonny.mancilla854@educacionbogota.edu.co
Doctorando en Educación
Instituto Pedagógico Rural " Gervasio Rubio" (IPRGR)
Venezuela

Rosa Palomino Salas
Orcid: 0009-0008-9011-1965
rosa.palomino.inforac@est.upel.edu.ve
Doctorando en Educación
Instituto Pedagógico Rural " Gervasio Rubio" (IPRGR)
Venezuela

Recibido: 08/01/2026 **Revisado: 12/02/2026** **Aprobado: 16/03/2026**

RESUMEN

Esta investigación compara ChatGPT y Gemini como asistentes para el aprendizaje inicial del desarrollo web, proporcionándoles como contexto una aplicación CRUD con arquitectura MVC (HTML, JavaScript, PHP y MySQL) entregada como proyecto funcional. Mediante un diseño mixto exploratorio, un evaluador experto analizó las respuestas a 12 preguntas estandarizadas en tres sesiones, valorando precisión técnica, claridad, profundidad pedagógica, utilidad experimental y consistencia. Gemini mostró una ligera ventaja en precisión técnica; ChatGPT destacó en profundidad pedagógica (con una diferencia del 5%). Ambas IAG presentaron alta consistencia y utilidad similar, lo que respalda su uso en tutorías personalizadas y justifica estudios empíricos futuros con estudiantes.

PALABRAS CLAVE: Inteligencia Artificial Generativa, Desarrollo Web, Asistencia Pedagógica

https://doi.org/10.56219/lneaimaginaria.v25i1.5858

Autora:
Yhonny Mancilla Díaz - Rosa Palomino Salas

EVALUATION OF CHATGPT AND GEMINI FOR BASIC WEB DEVELOPMENT LEARNING

ABSTRACT

This research compares ChatGPT and Gemini as learning aids for initial web development, using a functional CRUD application with an MVC architecture (HTML, JavaScript, PHP, and MySQL) as a context. Using an exploratory mixed-methods design, an expert evaluator analyzed responses to 12 standardized questions across three sessions, assessing technical accuracy, clarity, pedagogical depth, experimental usefulness, and consistency. Gemini showed a slight advantage in technical accuracy; ChatGPT stood out in pedagogical depth (with a 5% difference). Both learning aids demonstrated high consistency and similar usefulness, supporting their use in personalized tutoring and justifying future empirical studies with students.

Keywords. Generative Artificial Intelligence, Web Development, Pedagogical Support

https://doi.org/10.56219/lneaimaginaria.v25i1.5858

LÍNEA IMAGINARIA
REVISTA DE INVESTIGACIÓN EN ESTUDIOS SOCIALES Y DE FRONTERA

Número 25. Vol. 1 (2026), pp. 1- 17 / Edición Continua
ISSN en Línea: 2477-9415 / ISSN Impreso: 2477-9415
Depósito Legal: pp 200602TA2254

REPORTE DE INVESTIGACIÓN

INTRODUCCIÓN

Antecedentes de la investigación

En la actualidad, el desarrollo web se ha consolidado como una competencia esencial en la formación de profesionales en áreas tecnológicas y educativas. El uso de arquitecturas como Modelo Vista Controlador (MVC) (Ahmad et al., 2022; Anuar et al., 2022; De Petrillo & Mussbacher, 2024; Espinosa-Hurtado, 2021; Ponce Atencio et al., 2021) y de herramientas que integran lenguajes como HTML, JavaScript, PHP y MySQL (Ponce Atencio et al., 2021) permite a los estudiantes acercarse a entornos productivos reales, aunque con un nivel de complejidad que puede dificultar el aprendizaje inicial. En este escenario, se han propuesto soluciones basadas en la generación automática de código (Durai et al., 2022; Ponce Atencio et al., 2021; Vukosav et al., 2025), que buscan facilitar la comprensión de la lógica de programación y del funcionamiento de las aplicaciones web.

De manera complementaria, la integración de tecnologías de inteligencia artificial ha abierto nuevas posibilidades en los procesos de enseñanza y aprendizaje (Alpizar-Chacon & Keuning, 2025; Flórez Muñoz et al., 2024; Shanuka et al., 2024). Sistemas como ChatGPT o Gemini no solo ofrecen respuestas inmediatas, sino que también pueden adaptarse a distintos niveles de conocimiento (Shanuka et al., 2024), lo que contribuye a la personalización educativa (Husain, 2024). Sin embargo, la efectividad de dichas herramientas en contextos específicos, como la enseñanza del desarrollo de

aplicaciones CRUD(Create, Read, Update, Delete) con arquitectura MVC (Ahmad et al., 2021; Anuar et al., 2022; De Petrillo & Mussbacher, 2024), requiere ser evaluada con criterios que consideren tanto la pertinencia técnica como el valor pedagógico de las interacciones generadas (Alpizar-Chacon & Keuning, 2025).

Por lo anterior, se hace necesario un estudio que articule las posibilidades de los generadores de código con la mediación de inteligencias artificiales conversacionales. Este enfoque empírico permite situar la investigación en un campo donde confluyen los desarrollos tecnológicos con las prácticas educativas (Alpizar-Chacon & Keuning, 2025; De Petrillo & Mussbacher, 2024), favoreciendo la identificación de patrones de desempeño, fortalezas y limitaciones de las IA evaluadas (Flórez Muñoz et al., 2024; Shanuka et al., 2024). Así, el análisis busca aportar a la discusión académica sobre la pertinencia de estas herramientas en la formación inicial en desarrollo web y en la creación de apoyos didácticos para los docentes.

https://doi.org/10.56219/lneaimaginaria.v25i1.5858

Número 25. Vol. 1 (2026), pp. 1- 17 / Edición Continua
ISSN en Línea: 2477-9415 / ISSN Impreso: 2477-9415
Depósito Legal: pp 200602TA2254

Pregunta de investigación

¿Cuál de los modelos de inteligencia artificial, ChatGPT o Gemini, demuestra un mejor desempeño para ser utilizado como asistente de aprendizaje para estudiantes principiantes y de apoyo para profesores en el desarrollo web con tecnologías como HTML, JavaScript, PHP y MySQL?. Para responder a esta pregunta, el estudio evaluó la capacidad de ambas IAs para comprender y trabajar con una base de código fuente generada automáticamente, a fin de determinar su efectividad en la resolución de problemas, explicación de conceptos y generación de código adicional o modificaciones.

Objetivos de la investigación

Objetivo general

Evaluar el desempeño de ChatGPT y Gemini como asistentes de aprendizaje en la formación inicial en desarrollo web.

Objetivos específicos

- Diseñar e implementar un generador automático de código fuente para aplicaciones web funcionales tipo CRUD, basado en la arquitectura MVC con HTML, JavaScript, PHP y MySQL, que produzca un archivo .zip, a fin de proporcionar un contexto realista y funcional para la interacción con sistemas de inteligencia artificial como ChatGPT y Gemini.

doi

- Evaluar la efectividad de sistemas de inteligencia artificial, como ChatGPT y Gemini, en la asistencia al aprendizaje del desarrollo web, a partir de un contexto experimental. Para ello, se establecerá un conjunto de preguntas que serán respondidas por cada IA y analizadas por un evaluador experto.

Justificación de la investigación

La investigación se justifica en la necesidad de fortalecer los procesos de enseñanza y aprendizaje en el área del desarrollo web, un campo que demanda competencias técnicas cada vez más especializadas (Espinosa-Hurtado, 2021). En este sentido, la incorporación de herramientas de inteligencia artificial como asistentes pedagógicos representa una oportunidad para personalizar la experiencia educativa y optimizar la interacción con contenidos específicos, facilitando la comprensión de nuevos lenguajes de programación. Además, el uso de un generador automático de código CRUD funcional (Vukosav et al., 2025; Anuar et al., 2022) establece un escenario realista que simula los desafíos propios de la práctica profesional (De Petrillo & Mussbacher, 2024).

Asimismo, evaluar comparativamente el desempeño de diferentes sistemas de IA permite determinar con mayor objetividad cuál ofrece respuestas más pertinentes, claras y útiles en un contexto académico (Flórez Muñoz et al., 2024; Shanuka et al., 2024). Este tipo de análisis puede aportar al conocimiento sobre el papel de la IA en la mediación didáctica (Alpizar-Chacon & Keuning, 2025), y contribuir a generar criterios

https://doi.org/10.56219/lneaimaginaria.v25i1.5858

Número 25. Vol. 1 (2026), pp. 1- 17 / Edición Continua
ISSN en Línea: 2477-9415 / ISSN Impreso: 2477-9415
Depósito Legal: pp 200602TA2254

de selección informados para docentes e instituciones. El objetivo es potenciar el aprendizaje autónomo y facilitar el trabajo pedagógico, especialmente en entornos formativos con recursos limitados (Shanuka et al., 2024).

Finalmente, el impacto de la investigación radica en la posibilidad de recomendar, con base en evidencia, la integración de una IA específica como apoyo en la formación inicial de estudiantes de programación y en la labor de los docentes. Esto permitirá ampliar las estrategias de enseñanza, fomentar la inclusión digital y reducir las brechas de acceso al conocimiento. En consecuencia, se aportará al diseño de prácticas pedagógicas más innovadoras y eficientes, alineadas con las demandas actuales de la educación tecnológica.

El estudio se desarrolló en un contexto experimental y comparativo, evaluando el desempeño de ChatGPT y Gemini frente a doce preguntas sobre desarrollo web. Las respuestas fueron analizadas mediante rúbricas que midieron precisión técnica, claridad y coherencia conceptual, entre otros criterios. El proceso se realizó durante tres días consecutivos, lo que permitió obtener y valorar de forma controlada los resultados generados por ambas IAG.

El presente artículo se organiza en las secciones marco teórico, metodología, resultados, conclusiones, agradecimientos, referencias y anexos. En particular, en el apartado de metodología se describe el diseño de la investigación, centrándose en el desarrollo e implementación de un generador de código utilizado para crear

aplicaciones web básicas empaquetadas en archivos .zip. Estas aplicaciones se entregan como contexto base a las IAG con el fin de evaluar su desempeño a partir de sus respuestas a 12 preguntas.

MARCO TEORICO

Fundamentación teórica

Las tecnologías de IA aplicada a la educación han demostrado el potencial competitivo de los grandes modelos de lenguaje y ChatGPT en la enseñanza de la programación. Según (Quevedo et al., 2025) concluyeron en una revisión sistemática que “los modelos de lenguaje mejoran la productividad y motivación de los estudiantes en la generación de código, aunque requieren formación docente adecuada para evitar dependencia o mal uso” (p. 61). Esto ilustra el uso de IA generativa incluso más allá de aumentar la velocidad de escritura de código para proporcionar un contexto adecuado para el aprendizaje autodirigido.

Por otro lado, los autores enfatizan la necesidad de alguna forma de apoyo pedagógico que fomente la verificación crítica y la reflexión para que los educandos entiendan el proceso en lugar de simplemente replicar soluciones. Esta revisión de literatura justifica la relevancia de la comparación entre ChatGPT y Gemini como asistentes de aprendizaje, centrándose en las ventajas y desventajas en entornos educativos.

https://doi.org/10.56219/lneaimaginaria.v25i1.5858

LÍNEA IMAGINARIA
REVISTA DE INVESTIGACIÓN EN ESTUDIOS SOCIALES Y DE FRONTERA

Número 25. Vol. 1 (2026), pp. 1- 17 / Edición Continua
ISSN en Línea: 2477-9415 / ISSN Impreso: 2477-9415
Depósito Legal: pp 200602TA2254

REPORTE DE INVESTIGACIÓN

Es importante resaltar, que el 'Model View Controller' o 'MVC' es quizás uno de los más importantes. Aprende a utilizar el Modelo Controlador Vista en el desarrollo del framework web PHP. Castillo Yagual y Coronel Suárez (2023) sugieren que "los frameworks PHP basados en MVC permiten modular el código, simplificar operaciones CRUD, y permiten, en mejor medida, entender la estructura de una aplicación web" (p. 72). En este caso, los principales beneficiados son los disciplinantes, ya que el sistema les permite diferenciar los elementos que lo conforman, lo que como resultado tiene una mejor organización y escalabilidad del software.

Asimismo, el uso de MVC, especialmente con IA educativa, puede mejorar el aprendizaje de temas muy complejos, ya que las herramientas generativas pueden 'anidar' al aprendiz en el proceso de definir modelos, controlador y vista. Así, el marco teórico se posiciona en la intersección de la teoría del aprendizaje y la investigación reciente que enfatiza la utilidad de MVC y los modelos de framework de lenguaje como tecnología pedagógica en la educación digital.

Ahora bien, el uso de herramientas de inteligencia artificial en entornos educativos está en consonancia con las teorías de aprendizaje constructivista y socio constructivista que postulan que el aprendiz asume un papel activo en la adquisición de conocimientos y se beneficia de retroalimentación inmediata y personalizada. Más específicamente, la IA puede servir como mediador y facilitador de tareas individualizadas adaptadas al aprendiz, dentro de la zona de desarrollo

próximo Vygotsky (1979) afirmó que “la zona de desarrollo próximo es la distancia entre el nivel de desarrollo real y el nivel de desarrollo potencial” (p. 133). Tal idea sugiere que el estudiante alcanzará logros más altos con mediadores guiando su proceso, ya sean maestros, compañeros o tecnología.

Marco conceptual

La inteligencia artificial educativa se entiende como la aplicación de tecnología computacional destinada a asistir, complementar y mejorar los procesos de enseñanza y aprendizaje. Para Bustamante Bula y Camacho Bonilla (2024) “La IA en las escuelas mejora la calidad del aprendizaje y optimiza los procesos educativos, aunque requiere formación docente y marcos éticos” (p. 75). Dentro de esta área, los sistemas de modelos de lenguaje grandes (LLM), como ChatGPT y Gemini, son capaces de generar y comprender lenguaje natural, lo que les permite servir como posibles ayudas para los estudiantes de nivel básico que realizan ejercicios técnicos.

Estos instrumentos ayudan en la resolución de consultas, generación de código de ejemplo y provisión de retroalimentación instantánea, los cuales son críticos para el aprendizaje autodirigido. Desde una perspectiva educativa, los LLM son marcos en la inteligencia artificial que los teóricos del aprendizaje describir como agentes pedagógicos. Ellos trasladan la interacción alumno-tecnología hacia un aprendizaje más individualizado y autodirigido.

https://doi.org/10.56219/lneaimaginaria.v25i1.5858

LÍNEA IMAGINARIA
REVISTA DE INVESTIGACIÓN EN ESTUDIOS SOCIALES Y DE FRONTERA

Número 25. Vol. 1 (2026), pp. 1- 17 / Edición Continua
ISSN en Línea: 2477-9415 / ISSN Impreso: 2477-9415
Depósito Legal: pp 200602TA2254

REPORTE DE INVESTIGACIÓN

De igual forma, la investigación sobre la educación de inteligencia artificial ha propuesto marcos conceptuales para guiar la integración de modelos generativos en los procesos de enseñanza. Según (Anchundia et al., 2024), “la personalización de aprendizaje involucra el dominio cognitivo convirtiendo información en conocimiento mediante ejercicios de pensamiento crítico y creativo” (p. 4), esto refleja la capacidad de la IA para adaptar el contenido al nivel del alumno. Desde un punto de vista metodológico, las estrategias de adaptación de contenido con IA se proponen como mecanismos efectivos para abordar las diferencias individuales, mejorar la autorregulación y fomentar la comprensión de conceptos técnicos.

En este sentido, los modelos de lenguaje no solo generan respuestas, sino que pueden servir como asistentes pedagógicos para los estudiantes en la fase de inicio de la adquisición de habilidades digitales, siempre que su uso esté respaldado por un diseño educativo y la formación docente.

Por otra parte, El desarrollo web básico incluye el uso integrado de lenguajes de marcado, programación y bases de datos para construir aplicaciones dinámicas. Castillo Yagual y Coronel Suárez (2023) afirman que “la arquitectura Modelo-Vista-Controlador permite modularizar el código y simplificar las operaciones CRUD, lo que facilita el proceso de aprendizaje de los estudiantes” (p. 72). En este sentido, tecnologías como HTML, JavaScript, PHP y MySQL constituyen la base sobre la cual los estudiantes desarrollan habilidades introductorias en programación web.

La arquitectura MVC, al desacoplar los componentes de una aplicación, ayuda a los estudiantes que inician a entender mejor la interacción entre presentación, lógica y datos. Así, el concepto de MVC es crítico en este estudio, ya que sirve como el marco pedagógico dentro del cual se evalúa la efectividad de ChatGPT y Gemini como asistentes de aprendizaje en desarrollo web básico.

Bases legales

En Colombia la ley sobre la protección de datos personales llamada Ley 1581 de 2012. Establece que "la Constitución garantiza el derecho a conocer los datos recolectados sobre las personas en bases de datos y el derecho mencionado en el artículo 15 de la Constitución Política." (Ley 1581 de 2012. Art. 1). Esta regulación es de suma importancia para la investigación, para el despliegue de asistentes de IA como ChatGPT o Gemini, respecto a los datos procesados de los estudiantes del sistema que involucran datos sensibles que deben ser protegidos de acuerdo con la ley, el propósito, la seguridad, la confidencialidad y la autorización previa.

De igual forma, el Decreto 1377 de 2013 es una regulación parcial de la Ley 1581, de la cual lo más relevante es la autorización del titular, las políticas de tratamiento, la demostración de responsabilidad y las transferencias de datos del ámbito personal (Decreto 1377 de 2013, Artículos Iniciales). Así, cualquier sistema de

https://doi.org/10.56219/lneaimaginaria.v25i1.5858

LÍNEA IMAGINARIA
REVISTA DE INVESTIGACIÓN EN ESTUDIOS SOCIALES Y DE FRONTERA

Número 25. Vol. 1 (2026), pp. 1- 17 / Edición Continua
ISSN en Línea: 2477-9415 / ISSN Impreso: 2477-9415
Depósito Legal: pp 200602TA2254

REPORTE DE INVESTIGACIÓN

IA que registre, procese o difunda datos de los aprendices está sujeto al cumplimiento de estas regulaciones para ser legal.

Estudios y trabajos previos

El desarrollo de aplicaciones web a menudo implica tareas repetitivas como la implementación de operaciones CRUD (Crear, Leer, Actualizar, Eliminar), las cuales constituyen la base de la interacción con los datos (Anuar et al., 2022; Setyautami et al., 2025). Por ello, la generación automática de código ha surgido como una estrategia para optimizar este proceso, permitiendo a los desarrolladores enfocarse en la lógica de negocio más compleja en lugar de en la codificación de estructuras básicas (Durai et al., 2022; Vukosav et al., 2025). Investigaciones previas han explorado la creación de herramientas que, a partir de modelos de datos o especificaciones, generan código funcional para el backend y el frontend, frecuentemente bajo la arquitectura Modelo-Vista-Controlador (MVC), la cual promueve la organización y mantenibilidad del código (Ahmad et al., 2021; De Petrillo & Mussbacher, 2024).

Asimismo, el enfoque de Desarrollo Dirigido por Modelos (MDE) se ha consolidado como una metodología robusta que utiliza modelos abstractos, como diagramas UML, para generar implementaciones de software de manera automatizada (Ahmad et al., 2021; Vukosav et al., 2025). Estos enfoques buscan reducir el tiempo de

desarrollo y los errores humanos, aunque a menudo requieren una curva de aprendizaje para dominar las herramientas de modelado (Vukosav et al., 2025). La literatura existente describe generadores que transforman modelos de dominio o esquemas de bases de datos en aplicaciones web funcionales, demostrando la viabilidad de automatizar la creación de sistemas complejos y personalizables (Ponce Atencio et al., 2021; Setyautami et al., 2025).

Recientemente, la Inteligencia Artificial generativa ha transformado el paradigma del desarrollo de software, con herramientas como ChatGPT y GitHub Copilot que asisten en la generación, depuración y explicación de código (Flórez Muñoz et al., 2024; Shanuka et al., 2024). Estudios comparativos han evaluado la calidad y precisión del código producido por diversas plataformas de IA, revelando variaciones significativas en su desempeño y destacando que, si bien son prometedoras, aún requieren supervisión humana para garantizar la calidad y seguridad (Flórez Muñoz et al., 2024; Shanuka et al., 2024). En el contexto educativo, se ha analizado el uso de estas herramientas como apoyo para estudiantes, quienes reportan un aumento en la productividad y el aprendizaje, pero también expresan preocupación por la dependencia y la necesidad de aprender a formular prompts efectivos (Alpizar-Chacon & Keuning, 2025).

https://doi.org/10.56219/lneaimaginaria.v25i1.5858

LÍNEA IMAGINARIA
REVISTA DE INVESTIGACIÓN EN ESTUDIOS SOCIALES Y DE FRONTERA
Número 25. Vol. 1 (2026), pp. 1- 17 / Edición Continua
ISSN en Línea: 2477-9415 / ISSN Impreso: 2477-9415
Depósito Legal: pp 200602TA2254
REPORTE DE INVESTIGACIÓN

METODOLOGIA

Enfoque

Esta investigación se clasifica como exploratoria, ya que busca identificar el potencial de dos inteligencias artificiales generativas (IAGs), específicamente ChatGPT y Gemini, en el apoyo al aprendizaje de desarrollo web para estudiantes principiantes. Utilizando un generador de código CRUD con arquitectura MVC, que empaqueta aplicaciones web en un archivo ZIP como contexto base, se evalúa la efectividad de las respuestas de las IAGs a un conjunto de preguntas, con énfasis en su capacidad para acompañar procesos educativos. Por consiguiente, este nivel permite un análisis preliminar que oriente recomendaciones para profesores y alumnos, sin pretender generalizaciones amplias.

Además, el estudio incorpora elementos descriptivos al medir cualitativa y cuantitativamente las respuestas mediante métricas definidas por un evaluador experto, lo que facilita la comparación entre las IAGs. De esta manera, se describe el desempeño en términos de precisión, claridad y utilidad pedagógica, destacando posibles fortalezas en escenarios de enseñanza. Por lo tanto, esta aproximación descriptiva complementa el carácter exploratorio, proporcionando datos iniciales sobre la viabilidad de las IAGs como asistentes educativos.

Finalmente, aunque el enfoque no es explicativo, ya que no profundiza en causas subyacentes del rendimiento de las IAGs, sí establece bases para investigaciones futuras más avanzadas. Por lo tanto, este nivel de investigación se alinea con objetivos preliminares, priorizando la evaluación práctica sobre teorizaciones complejas. Así, contribuye a la metodología general al delimitar el alcance del estudio en el contexto de la educación en desarrollo web.

Diseño de la investigación

El diseño de la investigación corresponde a un enfoque de tipo comparativo-causal, dado que se busca analizar el desempeño de dos inteligencias artificiales frente a un mismo contexto de prueba. En este caso, una aplicación web básica empaquetada en un archivo .zip constituye el insumo inicial, mientras que las respuestas de los modelos se convierten en el objeto de análisis comparativo.

De esta forma, se contrasta el desempeño de las IA a partir de indicadores previamente definidos y observables, tales como precisión técnica, coherencia explicativa y pertinencia pedagógica, examinando las relaciones causales implícitas entre el contexto utilizado y la calidad de la respuesta, con el fin de obtener evidencia objetiva que oriente la toma de decisiones en escenarios educativos.

https://doi.org/10.56219/lneaimaginaria.v25i1.5858

LÍNEA IMAGINARIA
REVISTA DE INVESTIGACIÓN EN ESTUDIOS SOCIALES Y DE FRONTERA

Número 25. Vol. 1 (2026), pp. 1- 17 / Edición Continua
ISSN en Línea: 2477-9415 / ISSN Impreso: 2477-9415
Depósito Legal: pp 200602TA2254

REPORTE DE INVESTIGACIÓN

Las fases del estudio comprenden, en primer lugar, el diseño e implementación de un generador de aplicaciones web básicas de tipo CRUD. En segundo lugar, se elaboró una metodología para evaluar el desempeño de dos inteligencias artificiales, a las cuales se les suministra como contexto base un archivo comprimido (.zip) con la aplicación web generada. La tercera fase corresponde a la aplicación de dicha metodología, junto con el análisis y la presentación de los resultados.

Diseño e implementación de un generador de código de aplicación web

Se diseñó e implementó una herramienta que permite transformar una tabla de datos en una aplicación web básica de tipo CRUD, basada en la arquitectura MVC (Modelo Vista Controlador).

- **Flujo de uso del sistema.**

El flujo de uso del sistema se ilustra en la figura 1. El proceso comienza cuando el usuario prepara una tabla con datos y la copia desde una hoja de cálculo o desde cualquier archivo, quedando almacenada en el portapapeles de Windows. Posteriormente, al ejecutar el generador y presionar el botón Generar, entra en funcionamiento el componente denominado Parser. Este módulo procesa la tabla, identifica metadatos y datos relevantes como los nombres de los campos y sus

respectivos tipos, luego transforma dicha información en un objeto estructurado. Los tipos de datos soportados por el generador son tipo texto, entero, real, email y tipo fecha.

De manera previa, el administrador del sistema debe garantizar la existencia de un archivo de configuración, en el cual se especifiquen parámetros esenciales, tales como la conexión al servidor de base de datos MySQL y la configuración del servidor web local. Una vez cumplido este requisito, el Parser envía tanto la tabla procesada como la configuración al componente Generador. Este último se encarga de construir el código fuente de la aplicación y de producir los archivos correspondientes al FrontEnd y al BackEnd. Finalmente, el sistema crea una estructura de carpetas donde organiza los archivos generados y devuelve al usuario un paquete comprimido (.zip) que contiene la aplicación CRUD completamente funcional, lista para ejecutarse en la computadora local.

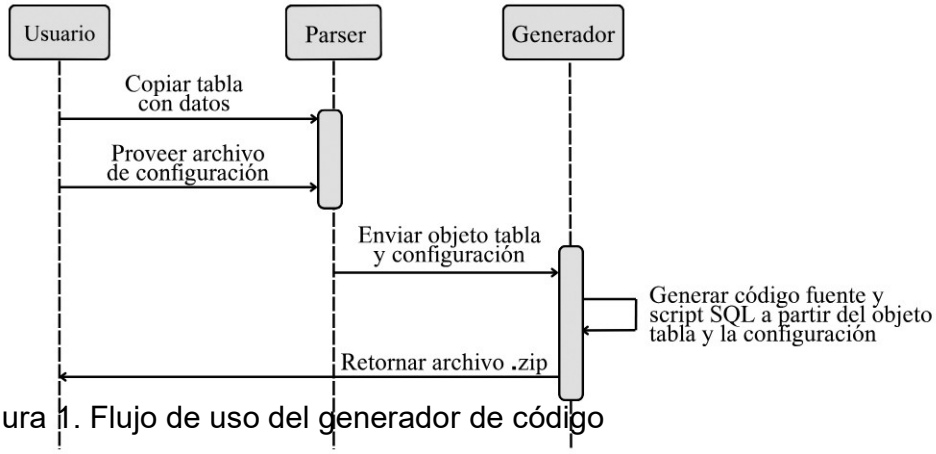


Figura 1. Flujo de uso del generador de código

https://doi.org/10.56219/lneaimaginaria.v25i1.5858

LÍNEA IMAGINARIA
REVISTA DE INVESTIGACIÓN EN ESTUDIOS SOCIALES Y DE FRONTERA
Número 25. Vol. 1 (2026), pp. 1- 17 / Edición Continua
ISSN en Línea: 2477-9415 / ISSN Impreso: 2477-9415
Depósito Legal: pp 200602TA2254
REPORTE DE INVESTIGACIÓN

- **Arquitectura de la aplicación web generada e integración con la IA**

En la figura 2 se presenta la arquitectura de la aplicación web generada. El componente Frontend está conformado por la vista, construida con código HTML y CSS, y por el controlador, implementado en JavaScript. Para la presentación de los datos se emplea el componente DataTables de jQuery, el cual facilita funciones como la paginación automática y la búsqueda de registros de manera ágil y práctica. Por su parte, el Backend está constituido por el modelo programado en PHP, estructurado como clase bajo el paradigma de orientación a objetos, que establece comunicación directa con una base de datos MySQL.

En cuanto a la integración con los usuarios y la inteligencia artificial, el proceso se desarrolla de forma directa. El usuario, tras estructurar una tabla de datos, la copia y la procesa mediante el generador, el cual construye automáticamente una aplicación web con la arquitectura previamente descrita y la empaqueta en un archivo comprimido (.zip). Este archivo se entrega como contexto base a la IA, con el propósito de ser utilizado como recurso en la primera fase de la investigación, específicamente para responder preguntas bajo la metodología de evaluación propuesta. A futuro, se espera que este esquema evolucione hacia un recurso didáctico, lo que permitirá su uso y nueva evaluación en entornos empíricos de enseñanza aprendizaje.

Autora:
Yhonny Mancilla Díaz - Rosa Palomino Salas

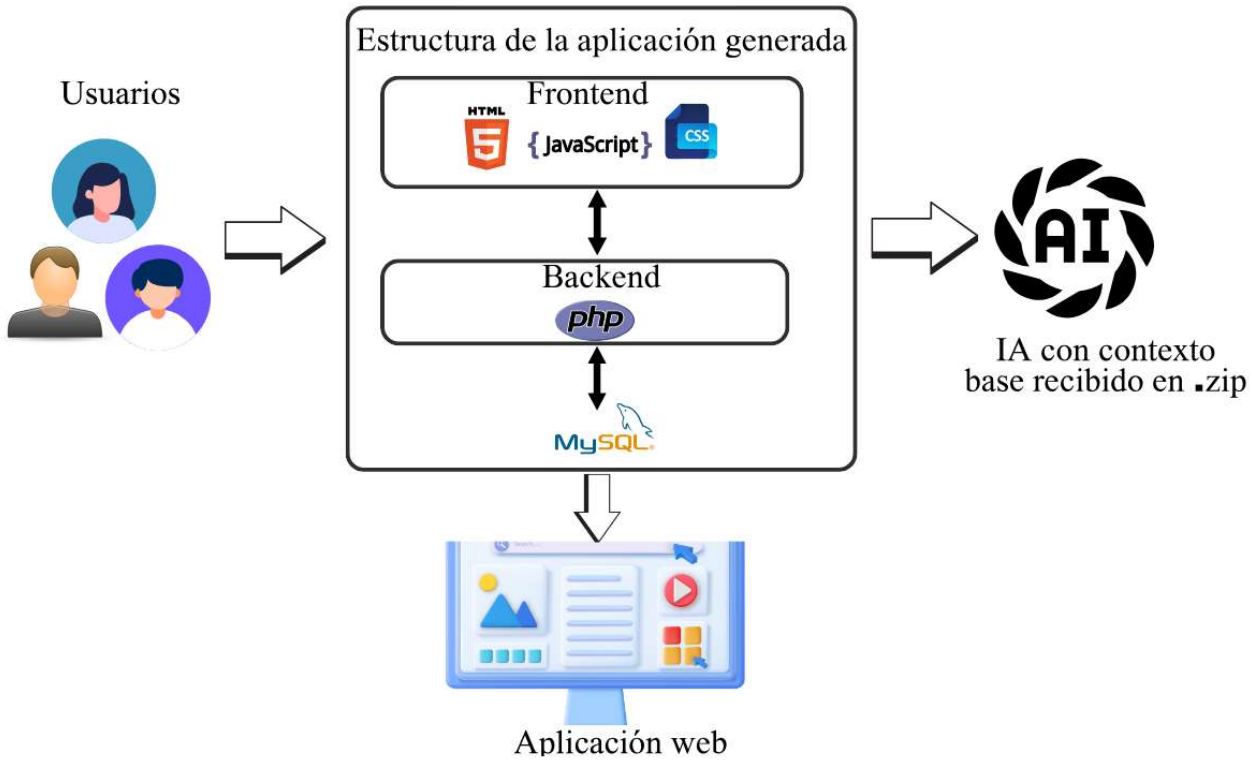


Figura 2. Arquitectura de la aplicación web generada e integración con la IA

- **Ejemplo de interfaz de aplicación web generada**

La imagen 3, contiene la interfaz de un ejemplo de aplicación web construida automáticamente por el generador de código, permitiendo las cuatro operaciones CRUD (create, read, update, delete). El caso de uso corresponde a una aplicación web para registrar contactos, con nombres, Email y Fecha de registro del contacto. Allí se

https://doi.org/10.56219/lneaimaginaria.v25i1.5858

LÍNEA IMAGINARIA
REVISTA DE INVESTIGACIÓN EN ESTUDIOS SOCIALES Y DE FRONTERA

Número 25. Vol. 1 (2026), pp. 1- 17 / Edición Continua
ISSN en Línea: 2477-9415 / ISSN Impreso: 2477-9415
Depósito Legal: pp 200602TA2254

REPORTE DE INVESTIGACIÓN

pueden visualizar los botones para eliminar y agregar un nuevo registro. La operación de edición se activa directamente en cada celda, cuando el usuario hace clic en ella.

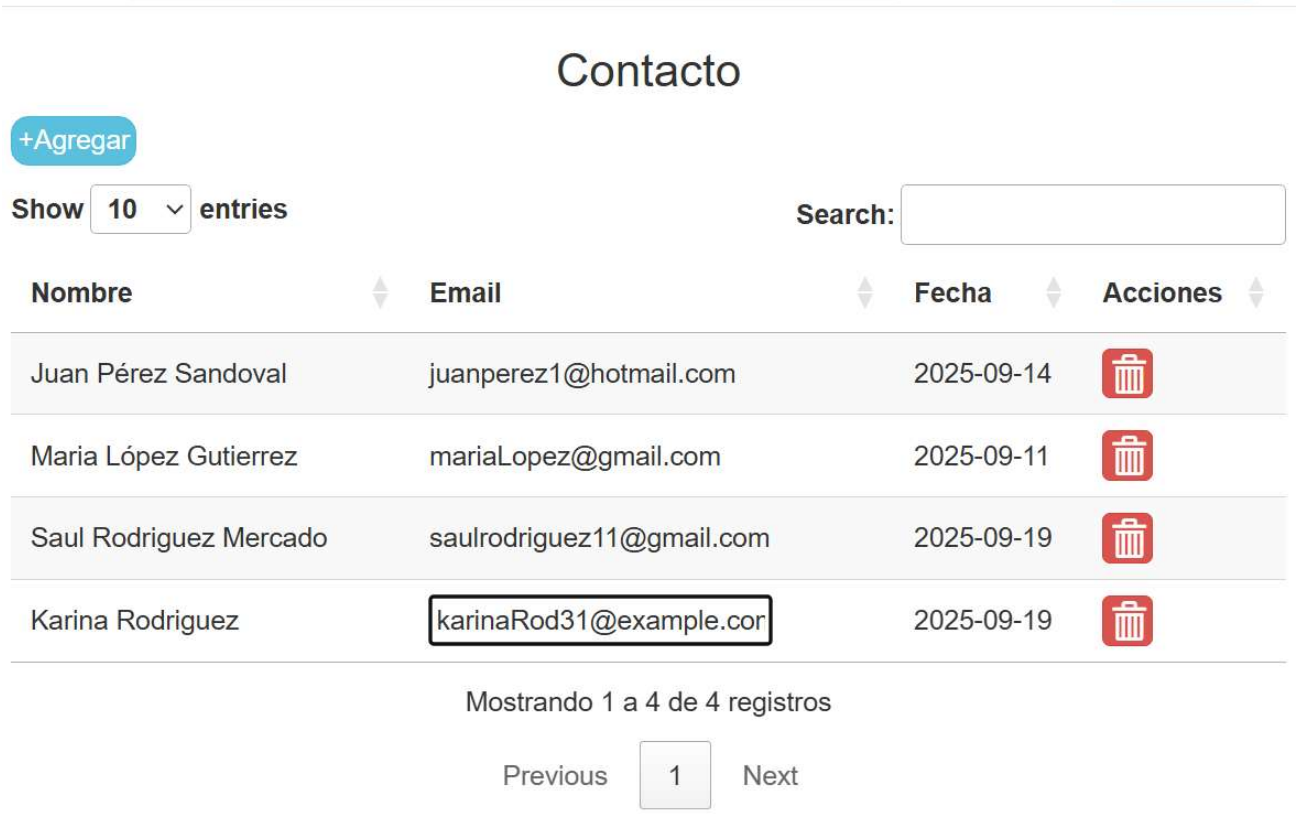


Figura 3. Ejemplo de aplicación web básica construida por el generador de código.

Unidades de análisis

En la presente investigación, la población se define como el conjunto de inteligencias artificiales conversacionales seleccionadas que son de acceso público y

que pueden procesar y generar código fuente, con un enfoque particular en aquellas optimizadas para tareas de desarrollo web. Este universo de IAs incluye, pero no se limita a, modelos ampliamente reconocidos como ChatGPT y Gemini, los cuales son los sujetos de investigación seleccionados para este estudio. Al ser un estudio previo con un número limitado de IAs, no se considera un muestreo probabilístico, sino que se opta por una selección intencional de estas dos herramientas debido a su popularidad y relevancia en el ámbito educativo y de desarrollo.

El método de selección se alinea con un enfoque no probabilístico de tipo intencional o por juicio, donde la elección de ChatGPT y Gemini se basa en criterios específicos y no en la aleatoriedad. Entre los criterios de inclusión para los sujetos de estudio se tienen ser una IA conversacional de fácil acceso, capaz de interpretar y generar código en los lenguajes específicos de la arquitectura MVC (HTML, Javascript, PHP) y ser pertinentemente conocida dentro de la comunidad de desarrolladores y educadores. Los criterios de exclusión abarcan a aquellas IAs que no cumplen con estas especificaciones, como modelos especializados en otras áreas o que requieran de configuraciones complejas que dificulten su uso por parte de estudiantes principiantes.

https://doi.org/10.56219/lneaimaginaria.v25i1.5858

LÍNEA IMAGINARIA
REVISTA DE INVESTIGACIÓN EN ESTUDIOS SOCIALES Y DE FRONTERA

Número 25. Vol. 1 (2026), pp. 1- 17 / Edición Continua
ISSN en Línea: 2477-9415 / ISSN Impreso: 2477-9415
Depósito Legal: pp 200602TA2254

REPORTE DE INVESTIGACIÓN

Técnicas de recolección de la información

Para la evaluación del desempeño de los modelos de IA, se adoptará una técnica experimental de carácter comparativo. Como instrumento principal se empleará un generador de código fuente automatizado, capaz de producir aplicaciones web CRUD funcionales bajo arquitectura MVC. Este sistema genera un archivo comprimido en formato .zip que contiene el código completo en HTML, JavaScript, PHP y MySQL, e integra la librería DataTables de jQuery. Dicho archivo funcionará como contexto base uniforme para todas las interacciones.

A continuación, este paquete de código se suministrará de manera idéntica a distintas IAs, como ChatGPT y Gemini, con el fin de garantizar condiciones homogéneas de análisis. Paralelamente, se aplicará un conjunto estandarizado de preguntas técnicas relacionadas con el código generado, que servirá como instrumento de recolección de datos y será administrado consistentemente a cada modelo. Las respuestas obtenidas se documentarán y archivarán de forma sistemática para su posterior revisión.

Finalmente, un evaluador experto aplicará una rúbrica compuesta por métricas mixtas, tanto cualitativas como cuantitativas, destinada a valorar la precisión, claridad y utilidad pedagógica de las respuestas. De este modo, se podrá contrastar de manera objetiva la efectividad de cada IA como recurso de apoyo en procesos de enseñanza-aprendizaje.

Metodología y proceso de análisis para evaluar la efectividad de las IAG en aprendizaje de desarrollo web básico

Dado el contexto de usar un archivo ZIP con código CRUD funcional (PHP, MySQL, JavaScript, HTML) para entregarla a una IA y simular el proceso de aprendizaje de un estudiante haciendo preguntas para entender, modificar y experimentar con el software, se plantea una metodología de tipo mixto basada en un evaluador experto. La metodología se divide en fases secuenciales, con énfasis en la objetividad del experto para minimizar sesgos.

La metodología propuesta consta de las fases de preparación, simulación de interacción, evaluación experta y análisis final. Inicialmente, se estructura un ejemplo de uso en una tabla de datos, luego se genera una aplicación web en un archivo .zip y se definen 12 preguntas estandarizadas sobre comprensión, arquitectura, modificaciones y experimentación, las cuales se aplican en sesiones interactivas donde un evaluador experto registra las respuestas de cada IAG para garantizar trazabilidad.

Posteriormente, califica las interacciones con métricas predefinidas, compara resultados entre IAGs, e implementa sugerencias en el código para verificar funcionalidad. Finalmente, se realiza un análisis y reporte con tablas que presentan promedios, y comparaciones, utilizando métricas cuantitativas y cualitativas estandarizadas, las cuáles se presentan en la tabla 1.

https://doi.org/10.56219/lneaimaginaria.v25i1.5858

LÍNEA IMAGINARIA
REVISTA DE INVESTIGACIÓN EN ESTUDIOS SOCIALES Y DE FRONTERA

Número 25. Vol. 1 (2026), pp. 1- 17 / Edición Continua
ISSN en Línea: 2477-9415 / ISSN Impreso: 2477-9415
Depósito Legal: pp 200602TA2254

REPORTE DE INVESTIGACIÓN

Métrica	Descripción	Tipo	Escala/Medida	Comparación
Precisión Técnica	Porcentaje de respuestas correctas sobre el código (e.g., explicación de arquitectura o modificaciones sin errores).	Cuantitativa	0-100% (basado en verificación experta).	Promedio por IAG; diferencia absoluta entre IAGs (e.g., ChatGPT: 85% vs. Gemini: 92%).
Claridad y Comprensibilidad	Nivel de explicación accesible para un principiante en desarrollo web.	Cualitativa	Escala Likert 1-5 (1: confusa, 5: clara y pedagógica).	Promedio y SD por categoría de preguntas; gráfico de barras comparativo.
Profundidad Pedagógica	Cobertura de conceptos (e.g., incluye ejemplos prácticos, prevención de errores).	Cualitativa	Escala Likert 1-5 (1: superficial, 5: profunda con guías paso a paso).	Promedio por IAG; correlación con éxito en modificaciones simuladas.
Utilidad para Experimentación	Efectividad en guiar modificaciones reales (e.g., % de sugerencias implementables sin fallos).	Cuantitativa	0-100% (basado en pruebas expertas de código sugerido).	Tasa de éxito por IAG; tabla de fallos comunes.
Consistencia Interactiva	Variabilidad en respuestas a preguntas repetidas o similares.	Cuantitativa	Coeficiente de variación (CV) en puntuaciones (bajo CV = alta consistencia).	CV promedio; umbral <20% para *efectiva*.

Tabla 1. Métricas a utilizar

A continuación, están las 12 preguntas referenciadas:

1. ¿Cuál es la estructura básica de la arquitectura MVC utilizada en el código y cómo se relacionan los componentes entre sí?
2. ¿Qué función cumple el archivo controlador.php en el flujo de datos entre la vista y el modelo?
3. ¿Cómo se manejan los errores de conexión a la base de datos en la clase UsuarioModel y qué tipo de excepciones se capturan?
4. Explica de manera sencilla cómo se carga y muestra la tabla de usuarios en la vista usando DataTables.
5. ¿Cómo se activa la edición de una celda en la tabla y qué pasa cuando el usuario presiona la tecla Enter?
6. Describe el proceso que sigue el sistema cuando un usuario hace clic en el botón "Agregar".
7. ¿Qué medidas de seguridad se implementan en el método actualizar del modelo para prevenir inyecciones SQL?

https://doi.org/10.56219/lneaimaginaria.v25i1.5858

LÍNEA IMAGINARIA
REVISTA DE INVESTIGACIÓN EN ESTUDIOS SOCIALES Y DE FRONTERA

Número 25. Vol. 1 (2026), pp. 1- 17 / Edición Continua
ISSN en Línea: 2477-9415 / ISSN Impreso: 2477-9415
Depósito Legal: pp 200602TA2254

REPORTE DE INVESTIGACIÓN

8. ¿Cómo se podría extender el código para validar que el email tenga un formato correcto
   antes de guardarlo?
9. Propón un ejemplo de cómo se podría agregar un campo adicional (por ejemplo, "teléfono") al formulario de usuario, incluyendo cambios en la base de datos, modelo, controlador y vista.
10. Modifica el código para que la columna "Fecha" muestre el formato dd/mm/aaaa en lugar de aaaa-mm-dd.
11. Implementa una validación en el cliente que impida eliminar un registro sin confirmación mediante un cuadro de diálogo personalizado (no usar confirm() nativo).
12. Añade un nuevo campo numérico "Edad" en la tabla, que sea editable directamente en la vista y se guarde en la base de datos.

Autora:
Yhonny Mancilla Díaz - Rosa Palomino Salas

RESULTADOS

Se realizó una evaluación comparativa entre las dos Inteligencias Artificiales Generativas (IAGs), Gemini y ChatGPT. Los resultados se obtuvieron a partir de respuestas a 12 preguntas sobre una aplicación web CRUD basada en arquitectura MVC. Estas preguntas se repitieron durante tres días consecutivos (25, 26 y 27 de septiembre de 2025), evaluando 5 métricas clave: Precisión técnica, Claridad y comprensibilidad, Profundidad pedagógica, Utilidad para experimentación y Consistencia interactiva. Las primeras cuatro métricas se aplicaron a tres preguntas específicas, utilizando rúbricas detalladas. La métrica de Consistencia interactiva se aplicó para identificar la variabilidad en respuestas a las mismas preguntas.

Los puntajes se normalizaron a una escala porcentual (0-1 para Precisión Técnica y Utilidad para Experimentación; conversión de Likert 1-5 a porcentaje para las otras dos). La evaluación se basó en un contexto realista proporcionado por un archivo ZIP generado automáticamente, que incluye código en HTML, CSS, JavaScript, PHP y MySQL para una tabla relativa a contactos o usuarios.

La Tabla 2 resume las puntuaciones promedio por métrica y día, derivadas de las evaluaciones detalladas por pregunta. En general, Gemini mostró una ligera superioridad en Precisión Técnica a lo largo de los tres días, con valores que oscilaron entre 0.93 y 0.95, mientras que ChatGPT se mantuvo entre 0.90 y 0.93. Para Claridad y Comprensibilidad, ambas IAGs obtuvieron puntuaciones altas, con ChatGPT

https://doi.org/10.56219/lneaimaginaria.v25i1.5858

LÍNEA IMAGINARIA
REVISTA DE INVESTIGACIÓN EN ESTUDIOS SOCIALES Y DE FRONTERA
Número 25. Vol. 1 (2026), pp. 1- 17 / Edición Continua
ISSN en Línea: 2477-9415 / ISSN Impreso: 2477-9415
Depósito Legal: pp 200602TA2254
REPORTE DE INVESTIGACIÓN

alcanzando 0.92 en el tercer día. En términos de profundidad pedagógica, ChatGPT superó a Gemini por un margen del 5%, con promedios de 0.86 y 0.81, respectivamente, reflejando una mayor consistencia en las explicaciones detalladas. Finalmente, en Utilidad para Experimentación, las puntuaciones fueron similares, rondando el 0.90-0.93, sin diferencias estadísticamente significativas (diferencia media < 0.03).

Día	Métrica	Puntuación Gemini	Puntuación ChatGPT
25 de septiembre	Precisión Técnica	0.93	0.90
	Claridad y Comprensibilidad	0.75	0.75
	Profundidad Pedagógica	0.75	0.83
	Utilidad para Experimentación	0.90	0.90
26 de septiembre	Precisión Técnica	0.95	0.92
	Claridad y Comprensibilidad	0.83	0.83
	Profundidad Pedagógica	0.75	0.83
	Utilidad para Experimentación	0.93	0.92
27 de septiembre	Precisión Técnica	0.95	0.93
	Claridad y Comprensibilidad	0.83	0.92
	Profundidad Pedagógica	0.92	0.92
	Utilidad para Experimentación	0.92	0.92

Tabla 2. Puntuaciones promedio por métrica y día para Gemini y ChatGPT

Las evaluaciones detalladas por pregunta revelaron patrones específicos. Por ejemplo, en Precisión Técnica (aplicada a preguntas 1-3), Gemini obtuvo consistentemente puntuaciones superiores o iguales (e.g., 0.83-0.95 en días 2 y 3), destacando su alineación con el código real del ZIP proporcionado, como en la descripción del manejo de errores con PDOException. En contraste, ChatGPT mostró omisiones menores, como no mencionar validaciones en el controlador. Para Claridad y Comprensibilidad (preguntas 4-6), ambas IAGs utilizaron estructuras pedagógicas efectivas, como pasos numerados y analogías, resultando en puntuaciones medias de 4-4.7 en escala Likert (convertidas a 0.75-0.92).

En Profundidad Pedagógica (preguntas 7-9), las respuestas incluyeron ejemplos de código y conexiones conceptuales, con puntuaciones promedio de 4-4.7, donde Gemini y ChatGPT empataron en varios casos al proporcionar guías paso a paso para modificaciones como agregar campos (e.g., "teléfono" o "edad"). Finalmente, para Utilidad para Experimentación (preguntas 10-12), las sugerencias de código fueron mayoritariamente implementables, con puntuaciones de 0.9-0.93, aunque con errores menores como falta de validación de fechas en algunas respuestas. En el Gráfico 1, se representa el resultado anterior.

https://doi.org/10.56219/lneaimaginaria.v25i1.5858

LÍNEA IMAGINARIA
REVISTA DE INVESTIGACIÓN EN ESTUDIOS SOCIALES Y DE FRONTERA

Número 25. Vol. 1 (2026), pp. 1- 17 / Edición Continua
ISSN en Línea: 2477-9415 / ISSN Impreso: 2477-9415
Depósito Legal: pp 200602TA2254

REPORTE DE INVESTIGACIÓN

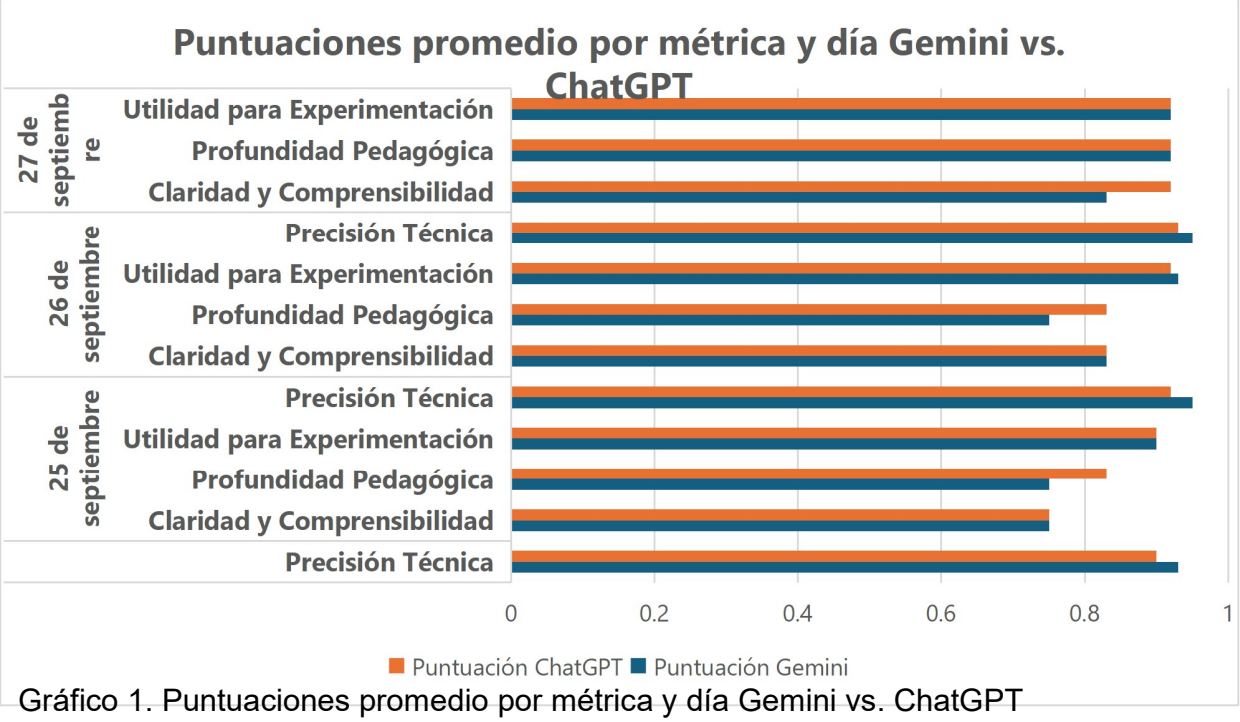


Gráfico 1. Puntuaciones promedio por métrica y día Gemini vs. ChatGPT

En el Gráfico 2 y Gráfico 3, ilustran la evolución diaria de las puntuaciones en las métricas evaluadas. Gemini mostró una tendencia ascendente en Profundidad Pedagógica (de 0.75 a 0.92), mientras que ChatGPT exhibió ascenso en Claridad y Comprensibilidad (de 0.75 a 0.92). No se observaron variaciones drásticas, con desviaciones estándar por métrica inferiores a 0.09, indicando consistencia en el rendimiento.

Autora:
Yhonny Mancilla Díaz - Rosa Palomino Salas

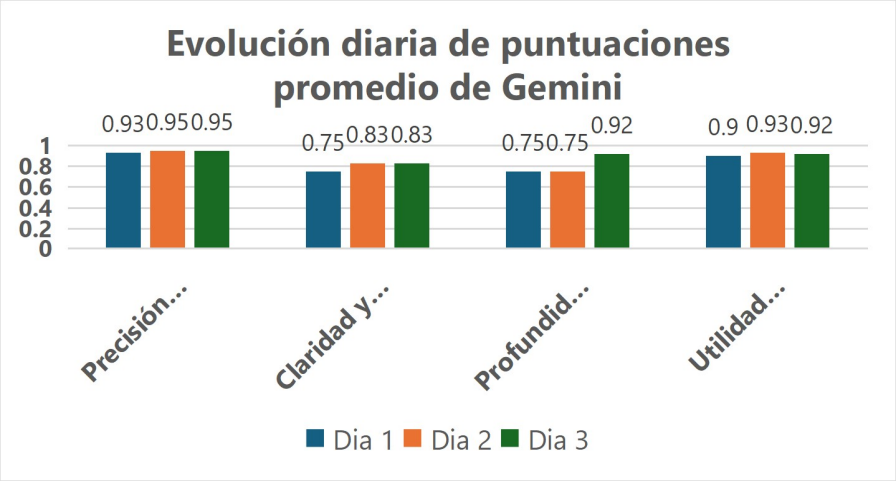
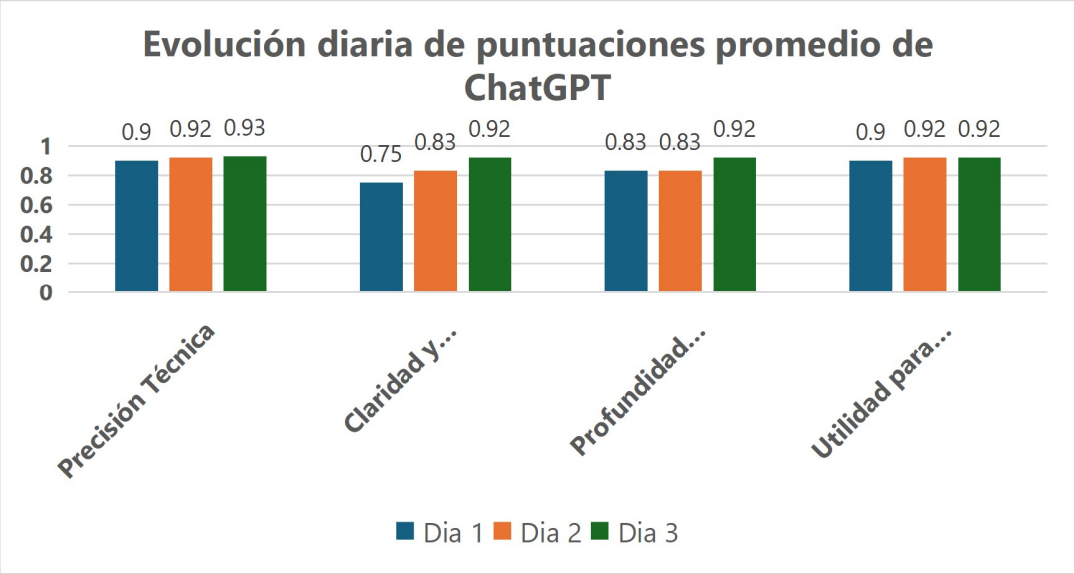


Gráfico 2. Evolución diaria de puntuaciones promedio por IAG Gemini



ráfico 3. Evolución diaria de puntuaciones promedio por IAG ChatGPT

https://doi.org/10.56219/lneaimaginaria.v25i1.5858

LÍNEA IMAGINARIA
REVISTA DE INVESTIGACIÓN EN ESTUDIOS SOCIALES Y DE FRONTERA

Número 25. Vol. 1 (2026), pp. 1- 17 / Edición Continua
ISSN en Línea: 2477-9415 / ISSN Impreso: 2477-9415
Depósito Legal: pp 200602TA2254

Para evaluar la utilidad de Gemini y ChatGPT como asistentes potenciales para estudiantes principiantes en el desarrollo de aplicaciones web básicas, se aplicó una rúbrica de consistencia interactiva basada en el coeficiente de variación (CV) a sus respuestas en 12 preguntas distribuidas a lo largo de tres días. Para Gemini, la mayoría de las métricas analizadas, como precisión técnica (CV=1.22%), claridad y comprensibilidad (CV=5.75%) y utilidad para experimentación (CV=1.67%), se clasificaron como muy consistentes, con variabilidad mínima que refleja respuestas casi idénticas en contenido y calidad; sin embargo, la profundidad pedagógica mostró un CV de 12.16%, lo que la posiciona como consistente, con variaciones mínimas pero predecibles en la esencia de las explicaciones. Esto sugiere que, aunque las respuestas mantienen una progresión temática estable, las leves fluctuaciones en profundidad podrían requerir una revisión ocasional por parte del usuario para asegurar uniformidad, resultando aun así adecuado para estudiantes principiantes que buscan fiabilidad en el aprendizaje inicial.

Por su parte, ChatGPT exhibió una consistencia más uniforme en general, con todos los aspectos, a saber, precisión técnica (CV=1.67%), claridad y comprensibilidad (CV=10.20%), profundidad pedagógica (CV=6.05%) y utilidad para experimentación (CV=1.26%), clasificándose como muy consistentes, gracias a una baja variabilidad que asegura respuestas similares con solo sinónimos o reformulaciones sutiles. En consecuencia, ambas inteligencias artificiales demuestran una estabilidad diaria

doi

notable, fomentando una experiencia educativa predecible; no obstante, ChatGPT podría representar una ventaja ligeramente más beneficiosa para principiantes al ofrecer mayor uniformidad en claridad y profundidad, lo que minimiza confusiones y facilita la experimentación práctica sin interrupciones en el flujo de aprendizaje.

Los hallazgos principales indican que tanto Gemini como ChatGPT demostraron un alto rendimiento en la mayoría de métricas evaluadas, con Gemini exhibiendo una ligera ventaja en Precisión Técnica y Utilidad para experimentación, especialmente en los días posteriores. Esto sugiere que Gemini podría ofrecer un mayor valor práctico para acompañar el aprendizaje de estudiantes principiantes en desarrollo web, al proporcionar explicaciones más alineadas con el código real y útiles para experimentación. Por ejemplo, en preguntas sobre manejo de errores (e.g., PDOException), Gemini integró detalles factuales con mayor exactitud, reduciendo omisiones que podrían confundir a principiantes.

Estos resultados se interpretan en el contexto de la capacidad de las IAGs para procesar un ZIP generado por un generador de código ágil, que crea aplicaciones CRUD MVC a partir de tablas de datos en el portapapeles. Este generador representa una innovación clave, ya que elimina barreras técnicas iniciales como la configuración manual de entornos, permitiendo a estudiantes con conocimientos básicos enfocarse en conceptos en lugar de sintaxis básica.

https://doi.org/10.56219/lneaimaginaria.v25i1.5858

LÍNEA IMAGINARIA
REVISTA DE INVESTIGACIÓN EN ESTUDIOS SOCIALES Y DE FRONTERA

Número 25. Vol. 1 (2026), pp. 1- 17 / Edición Continua
ISSN en Línea: 2477-9415 / ISSN Impreso: 2477-9415
Depósito Legal: pp 200602TA2254

Al generar código funcional en minutos, se reduce el tiempo de preparación (de horas a segundos) y adapta el aprendizaje a escenarios realistas, fomentando experimentación sin necesidad de configuraciones iniciales complejas o poco accesibles para principiantes. En comparación con métodos tradicionales (e.g., tutoriales estáticos), este enfoque tiene el potencial de acelerar el ciclo de aprendizaje, alineándose con principios de educación. No se encontró en la literatura existente un método similar para generar ZIPs de aplicaciones web básicas completas y funcionales de forma rápida a partir de una tabla copiada al portapapeles de windows, lo que posiciona esta herramienta como un avance para entornos educativos inclusivos.

Relacionando con la literatura revisada en los antecedentes, estudios como los de Vukosav et al. (2025) destacan cómo generadores de código bajan umbrales para novatos, similar a este caso donde el ZIP permite a las IAGs responder basadas en código realista, mejorando la relevancia pedagógica. Asimismo, investigaciones en IAGs educativas (e.g., Alpizar-Chacon & Keuning, 2025) enfatizan la importancia de claridad y profundidad. Para los resultados obtenidos en esta investigación, ambas IAGs superaron expectativas, pero ChatGPT se alineó mejor con rúbricas que valoran analogías y pasos ordenados de forma lógica.

doi

Sin embargo, el estudio presenta limitaciones. La evaluación se basó en solo tres días y 12 preguntas, potencialmente sesgada por variaciones diarias en las IAGs (e.g., actualizaciones de modelos). Además, no se incluyeron pruebas con estudiantes reales para validar la efectividad percibida.

Como conclusión de esta sección, los resultados destacan implicaciones para la integración de IAGs en educación; Gemini podría priorizarse para precisión en temas técnicos, mientras que el generador de código emerge como innovación para democratizar el aprendizaje, sugiriendo futuras direcciones como integrar IAGs con generadores en tiempo real para tutorías personalizadas. Investigaciones adicionales podrían explorar escalabilidad con más preguntas o métricas cualitativas.

https://doi.org/10.56219/lneaimaginaria.v25i1.5858

LÍNEA IMAGINARIA
REVISTA DE INVESTIGACIÓN EN ESTUDIOS SOCIALES Y DE FRONTERA

Número 25. Vol. 1 (2026), pp. 1- 17 / Edición Continua
ISSN en Línea: 2477-9415 / ISSN Impreso: 2477-9415
Depósito Legal: pp 200602TA2254

REPORTE DE INVESTIGACIÓN

CONCLUSIONES

Como conclusión, los resultados destacan implicaciones para la integración de IAGs en educación; Gemini podría priorizarse para precisión en temas técnicos, mientras que el generador de código emerge como innovación para democratizar el aprendizaje, sugiriendo futuras direcciones como integrar IAGs con generadores en tiempo real para tutorías personalizadas. Investigaciones adicionales podrían explorar escalabilidad con más preguntas o métricas cualitativas.

Los objetivos principales del estudio fueron comparar dos IAGs en su capacidad para apoyar el aprendizaje de estudiantes principiantes en desarrollo web, evaluando métricas clave a través de respuestas a preguntas sobre una aplicación CRUD MVC generada ágilmente. En relación con estos objetivos, los resultados más importantes revelan que Gemini superó ligeramente a ChatGPT en Precisión Técnica (promedio 0.94 vs. 0.92) y Utilidad para experimentación (0.92 vs. 0.91), mientras que ChatGPT superó en Profundidad Pedagógica (0.86 vs. 0.81) y Claridad y comprensibilidad (0.83 vs. 0.8) a Gemini. Esto indica que ambas son viables, pero ChatGPT ofrece una mayor precisión en la información y resulta más accesible para principiantes.

Las principales implicaciones incluyen el potencial de estas IAGs para transformar la educación en desarrollo web, especialmente cuando se combinan con innovaciones como el generador de código ágil, que elimina barreras técnicas y permite enfocarse en aprendizaje conceptual.

Para futuras investigaciones, se sugiere considerar los resultados obtenidos en este estudio preliminar para realizar evaluaciones empíricas de Gemini y ChatGPT en entornos reales de aprendizaje con estudiantes, ampliando así la validez de los hallazgos. Además, es esencial continuar la evaluación de otras herramientas de IAGs para determinar su capacidad de apoyo en el proceso de enseñanza aprendizaje del desarrollo web, integrando otras tecnologías. Por otra parte, para trabajos futuros que incorporen metodologías de generación ágil de aplicaciones, se recomienda incluir métricas robustas para evaluar la calidad del código generado. Finalmente, este tipo de estudios debería repetirse con una periodicidad razonable debido a la constante y acelerada evolución de los Modelos de Lenguaje Grandes (LLMs).

En síntesis, este estudio afirma que Gemini es más efectiva para acompañar experimentación en principiantes, mientras que una integración híbrida de ambas IAGs, apoyada por herramientas generadoras de código, podría revolucionar la educación en desarrollo web al hacerla más accesible y eficiente.

https://doi.org/10.56219/lneaimaginaria.v25i1.5858

LÍNEA IMAGINARIA
REVISTA DE INVESTIGACIÓN EN ESTUDIOS SOCIALES Y DE FRONTERA

Número 25. Vol. 1 (2026), pp. 1- 17 / Edición Continua
ISSN en Línea: 2477-9415 / ISSN Impreso: 2477-9415
Depósito Legal: pp 200602TA2254

REPORTE DE INVESTIGACIÓN

AGRADECIMIENTOS

Agradecemos de manera especial a la Dra. Yaritza Cova Jaime, cuyo acompañamiento y asesoría fueron fundamentales para la producción de este artículo de investigación.

REFERENCIAS

Alpizar-Chacon, I., & Keuning, H. (2025). Student's Use of Generative AI as a Support Tool in an Advanced Web Development Course. En Proceedings of... (pp. 1-7). ACM. [Disponible en línea: https://dl.acm.org/doi/pdf/10.1145/3724363.3729106]

Anuar, A. W., Kama, N., Azmi, A., & Rusli, H. M. (2022). Revisiting Web Application Development with Integrated Records Management Important Aspect using Re-CRUD. Journal of Information and Knowledge Management (JIKM), 12(1), 31–54. [Disponible en línea: https://ir.uitm.edu.my/id/eprint/65774/1/65774.pdf]

De Giusti, L. C., Villarreal, G. L., Ibañez, E. J., & De Giusti, A. E. (2023). Aprendizaje y enseñanza de programación: El desafío de herramientas de Inteligencia Artificial como ChatGPT. Repositorio Institucional de la Universidad Nacional de La Plata (SEDICI). [Disponible en línea: https://sedici.unlp.edu.ar/bitstream/handle/10915/165093/Documento_completo.pdf?sequence=1]

De Petrillo, E., & Mussbacher, G. (2024). FeatureLanguage: Automatic Generation of Application Backend for Model-Based Programming Course Projects. En Proceedings of MoDRE2024. [Disponible en línea: https://www.modre2024.ece.mcgill.ca/proceedings/MoDRE2024_1.pdf]

Espinosa-Hurtado, R. (2021). Análisis comparativo para la evaluación de frameworks usados en el desarrollo de aplicaciones web. CEDAMAZ, 11(2), 58–66. [Disponible en línea: https://revistas.unl.edu.ec/index.php/cedamaz/article/download/1182/853]

Flórez Muñoz, M. A., Jaramillo de la Torre, J. C., Pareja López, S., Herrera Sierra, S., & Candela-Uribe, C. A. (2024). Estudio comparativo de herramientas de generación de código por IA: evaluación de calidad y análisis de desempeño. Revista Ciencia e Ingeniería, 11(2), e12809577. [Disponible en línea: https://dialnet.unirioja.es/descarga/articulo/9685509.pdf]

Husain, A. (2024). Potentials of ChatGPT in Computer Programming: Insights from Programming Instructors. Journal of Information Technology Education: Research, pp. 23, 002. https://doi.org/10.28945/5240. [Disponible en línea: https://www.jite.org/documents/Vol23/JITE-Rv23Art002Husain9941.pdf]

Llanos Mosquera, J. M., Hidalgo Suarez, C. G., & Bucheli Guerrero, V. A. (2021). Una revisión sistemática sobre aula invertida y aprendizaje colaborativo apoyados en inteligencia artificial para el aprendizaje de programación. Tecnura, 25(69), 196–214. [Disponible en línea: http://www.scielo.org.co/pdf/tecn/v25n69/0123-921X-tecn-25-69-196.pdf]

Ponce Atencio, Y., Ibarra, M. J., Huanca Marin, J., & Holguin Holguin, E. (2021). Automatic Generation of Web Applications for Information Systems. Journal of

https://doi.org/10.56219/lneaimaginaria.v25i1.5858

Número 25. Vol. 1 (2026), pp. 1- 17 / Edición Continua
ISSN en Línea: 2477-9415 / ISSN Impreso: 2477-9415
Depósito Legal: pp 200602TA2254

Physics: Conference Series, 1860(1), 012019. [Disponible en línea: https://iopscience.iop.org/article/10.1088/1742-6596/1860/1/012019/pdf]

Shanuka, K. A. A., Wijayanayake, J., & Vidanage, K. (2024). Systematic literature review on analysing the impact of prompt engineering on efficiency, code quality, and security in CRUD application development. The Journal of Desk Research Review and Analysis, 2(1), 235–249. [Disponible en línea: https://jdrra.sljol.info/articles/57/files/679879ec570e2.pdf]

Vukosav, D., Banjac, D., Ljubojević, M., & Savić, M. (2025). CodeCrafter – Efficient Code Generator for Modern Single-page Web Applications. International Journal of Electrical Engineering and Computing, 9(1), 1–9. [Disponible en línea: https://ijeec.etf.ues.rs.ba/index.php/ijeec/article/download/196/127]

Durai, A. D., Ganesh, M., Mathew, R. M., & Anguraj, D. K. (2022). A novel approach with an extensive case study and experiment for automatic code generation from the XMI schema Of UML models. *The Journal of Supercomputing, 78*, 7677–7699. https://doi.org/10.1007/s11227-021-04164-x

Setyautami, M. R. A., Hähnle, R., Azurat, A., & Budiardjo, E. K. (2025). End-to-end development of product lines for web systems. *Software and Systems Modeling*. https://doi.org/10.1007/s10009-025-00784-3. [Disponible en línea: https://link.springer.com/content/pdf/10.1007/s10009-025-00784-3.pdf]

Ahmad, S. I., Rana, T., & Maqbool, A. (2021). A Model-Driven Framework for the Development of MVC-Based (Web) Application. *Arabian Journal for Science and Engineering, 47*, 1733–1747. 7https://doi.org/10.1007/s13369-021-06087-4.

Anchundia Parraga, J., Macías Zambrano, R. M., & Tubay Cevallos, L. A. (2024). La personalización del aprendizaje: estrategias de adaptación de contenido con inteligencia artificial en entornos educativos. Educación y Vínculos. Revista de Estudios Interdisciplinarios en Educación, (13). [Disponible en línea: https://pcient.uner.edu.ar/index.php/EyV/article/download/1940/2146]

Castillo Yagual, C. A., & Coronel Suárez, M. A. (2023). Frameworks PHP basados en la arquitectura Modelo-Vista-Controlador para desarrollo de aplicaciones web. Revista Científica y Tecnológica UPSE (RCTU), 10(1), 70–78. https://doi.org/10.26423/rctu.v10i1.703. [Disponible en línea: http://scielo.senescyt.gob.ec/pdf/rctu/v10n1/1390-7697-rctu-10-01-00070.pdf]

Bustamante Bula, R., & Camacho Bonilla, A. (2024). Inteligencia artificial (IA) en las escuelas: una revisión sistemática (2019-2023). Enunciación, 29(1), 62-82. https://doi.org/10.14483/22486798.22039. [Disponible en línea: http://www.scielo.org.co/pdf/enunc/v29n1/2248-6798-enunc-29-01-62.pdf]

Quevedo-Tumailli, V. F., Arias Calderón, S. E., Ortega Manjarrez, V. A., & Ortega-Tenezaca, B. (2025). Impacto de modelos de lenguaje de gran tamaño en calidad y eficiencia de generación de código: Revisión sistemática de literatura.

Novasinergia, 8(1), 52–66. https://doi.org/10.37135/ns.01.15.10. [Disponible en línea: http://scielo.senescyt.gob.ec/pdf/rns/v8n1/2631-2654-rns-8-01-00052.pdf]

Vygotsky, L. S. (1979). El desarrollo de los procesos psicológicos superiores. Barcelona: Editorial Crítica. [Disponible en línea: https://aulasvirtuales.udistrital.edu.co/pluginfile.php/943614/mod_resource/content/2/El-desarrollo-procesos.pdf]

Ley 1581 de 2012. Por la cual se dictan disposiciones generales para la protección de datos personales. [Disponible en línea: https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=49981]

Decreto 1377 de 2013. Por el cual se reglamenta parcialmente la Ley 1581 de 2012, en lo relacionado con autorización del titular, políticas y responsabilidad en el tratamiento de datos personales. República de Colombia. [Disponible en línea: https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=53646]

https://doi.org/10.56219/lneaimaginaria.v25i1.5858

LÍNEA IMAGINARIA
REVISTA DE INVESTIGACIÓN EN ESTUDIOS SOCIALES Y DE FRONTERA

Número 25. Vol. 1 (2026), pp. 1- 17 / Edición Continua
ISSN en Línea: 2477-9415 / ISSN Impreso: 2477-9415
Depósito Legal: pp 200602TA2254

REPORTE DE INVESTIGACIÓN

ANEXOS

Rúbricas utilizadas para la evaluación de las respuestas de las IAG

- **Rúbrica para precisión técnica**

Umbral de Puntuación	Criterios
0-20% (Muy Baja)	Más del 80% de hechos erróneos o inexactos (e.g., describe incorrectamente una query SQL como segura cuando no lo es).
21-40% (Baja)	60-80% de errores; algunos hechos correctos, pero con inexactitudes graves que invalidan la respuesta (e.g., confunde PHP con JavaScript en interacciones).
41-60% (Moderada)	40-60% de errores; mezcla hechos correctos con menores imprecisiones (e.g., explica arquitectura correctamente pero omite detalles menores).
61-80% (Alta)	20-40% de errores menores; la mayoría de hechos son precisos y verificables (e.g., describe modificaciones con solo un error sintáctico menor).
81-100% (Excelente)	Menos del 20% de errores; todos los hechos son factuales y alineados con el código real (e.g., explicación perfecta de flujo CRUD sin inexactitudes).

- **Rúbrica para Claridad y comprensibilidad**

Puntuación	Criterios
1 (Confusa)	Explicación desorganizada, con jerga técnica sin definición, oraciones complejas o incoherentes (e.g., "El CRUD implementa MVC mediante abstracciones relacionales").
2 (Poco Clara)	Algo de estructura, pero con jerga no explicada y oraciones largas que confunden (e.g., menciona "AJAX" sin aclarar qué es).

Autora:
Yhonny Mancilla Díaz - Rosa Palomino Salas

3 (Moderada)	Explicación básica y estructurada, pero con algunos términos no definidos o saltos lógicos (e.g., explica pasos, pero asume conocimiento previo mínimo).
4 (Clara)	Lenguaje simple, términos definidos, estructura lógica con viñetas o pasos numerados (e.g., "AJAX es una técnica para actualizar partes de la página sin recargar; aquí se usa para enviar datos al PHP").
5 (Muy Clara y Pedagógica)	Excepcionalmente accesible: usa analogías, repite conceptos clave, y adapta al nivel principiante con preguntas retóricas (e.g., "Imagina el HTML como la fachada de una casa; el PHP es el motor interno").

- **Rúbrica para Profundidad pedagógica**

Puntuación	**Criterios**
1 (Superficial)	Respuesta factual básica sin ejemplos ni contexto (e.g., "Esta función inserta datos").
2 (Básica)	Explica el concepto con un ejemplo mínimo, sin prevención de errores ni conexiones (e.g., "Agrega un input en HTML").
3 (Moderada)	Incluye un ejemplo práctico y menciona un error común, pero no ofrece pasos detallados (e.g., "Agrega input y actualiza query, cuidado con SQL injection").
4 (Avanzada)	Proporciona ejemplo práctico, prevención de errores y conexiones conceptuales (e.g., "Agrega input, usa prepared statements para evitar SQL injection; esto conecta con seguridad en MVC").
5 (Profunda)	Guía paso a paso con ejemplo funcional, prevención de errores, conexiones conceptuales y sugerencia de experimentación (e.g., "Paso 1: input HTML, Paso 2: query con prepared statements; prueba con var_dump() y observa errores").

https://doi.org/10.56219/lneaimaginaria.v25i1.5858

LÍNEA IMAGINARIA
REVISTA DE INVESTIGACIÓN EN ESTUDIOS SOCIALES Y DE FRONTERA

Número 25. Vol. 1 (2026), pp. 1- 17 / Edición Continua
ISSN en Línea: 2477-9415 / ISSN Impreso: 2477-9415
Depósito Legal: pp 200602TA2254

REPORTE DE INVESTIGACIÓN

- **Rúbrica para Utilidad para experimentación**

Umbral de Puntuación	Criterios
0-20% (Muy Baja)	Más del 80% de sugerencias fallan en implementación (e.g., código sugerido causa errores fatales).
21-40% (Baja)	60-80% fallos; sugerencias parciales, pero requieren correcciones mayores (e.g., modifica formulario pero ignora actualización DB).
41-60% (Moderada)	40-60% fallos; funciona en partes, con errores menores corregibles (e.g., sintaxis errónea pero lógica correcta).
61-80% (Alta)	20-40% fallos menores; la mayoría implementable con ajustes mínimos (e.g., código funciona tras un fix simple).
81-100% (Excelente)	Menos del 20% fallos; sugerencias completamente implementables y funcionales (e.g., guía que resulta en modificación exitosa sin errores).

- **Rúbrica para consistencia interactiva**

Umbral de CV (Coeficiente de variación)	Criterios
>50% (Muy Inconsistente)	Alta variabilidad; respuestas cambian drásticamente en precisión o profundidad (e.g., una vez explica bien, otra vez omite todo).
31-50% (Inconsistente)	Variabilidad moderada-alta; diferencias notables en detalles (e.g., varía en ejemplos proporcionados).
21-30% (Moderada)	Variabilidad aceptable; cambios menores pero consistentes en esencia (e.g., reformula pero mantiene hechos).
11-20% (Consistente)	Baja variabilidad; respuestas similares con variaciones mínimas (e.g., misma explicación con sinónimos).
0-10% (Muy Consistente)	Mínima variabilidad; respuestas casi idénticas en contenido y calidad.