



UNIVERSIDAD ESTATAL DE MILAGRO

REPÚBLICA DEL ECUADOR

**UNIVERSIDAD ESTATAL DE MILAGRO
FACULTAD DE POSGRADO**

VICERRECTORADO DE INVESTIGACIÓN Y POSGRADO

**ARTÍCULOS PROFESIONALES DE ALTO NIVEL
PREVIO A LA OBTENCIÓN DEL TÍTULO DE:**

MAGÍSTER EN INTELIGENCIA ARTIFICIAL PARA LA EDUCACIÓN

TEMA:

**Sistemas de tutoría inteligente basados en inteligencia artificial generativa:
una revisión sistemática según directrices PRISMA**

Autor:

Vergara Aguilera Edison Guillermo

Tutor:

Solís Granda Luis Eduardo

Milagro, 2025

Sistemas de tutoría inteligente basados en inteligencia artificial generativa: una revisión sistemática según directrices PRISMA

Intelligent tutoring systems based on generative artificial intelligence: a systematic review according to PRISMA guidelines

Autores:

Vergara Aguilera, Edison Guillermo

Universidad Estatal de Milagro | Facultad de Posgrados | Ecuador
evergaraa@unemi.edu.ec | ORCID: 0009-0007-9652-7446

Solis-Granda, Luis Eduardo

Universidad Estatal de Milagro | Facultad de Posgrados | Ecuador
lsolisg@unemi.edu.ec | ORCID: 0000-0002-1387-7333

RESUMEN

La integración de la inteligencia artificial generativa en la educación ha promovido el desarrollo de nueva tecnología centrada en mejorar la personalización del aprendizaje. En este sentido, los sistemas de tutoría inteligente han avanzado significativamente con la llegada de grandes modelos de lenguaje capaces de generar contenido educativo, ofrecer retroalimentación adaptativa y participar en interacciones conversacionales con los estudiantes. La presente investigación tiene como objetivo analizar los sistemas de tutoría inteligente basados en inteligencia artificial generativa mediante una revisión sistemática de la literatura. La metodología se desarrolló bajo el protocolo PRISMA, que permitió una organización rigurosa de las etapas de búsqueda, selección y análisis de los estudios. Se realizaron búsquedas en las bases de datos Web of Science, IEEE Xplore y SciELO mediante ecuaciones de búsqueda con operadores booleanos AND y OR. En la etapa de identificación se recopilieron 1 485 registros correspondientes al período 2020–2026, que se refinaron eliminando duplicados y aplicando criterios de inclusión y exclusión. Se seleccionaron nueve estudios tras la revisión de títulos, resúmenes y textos completos, siguiendo los criterios metodológicos establecidos. Los resultados evidencian que los sistemas de tutoría inteligente con inteligencia artificial generativa permiten personalizar el aprendizaje, crear materiales educativos de forma automática, ofrecer retroalimentación personalizada y mejorar las interacciones entre el estudiante y el sistema. La incorporación de modelos de lenguaje de gran escala y arquitecturas generativas como RAG y sistemas multiagente mejora la interacción y la adaptabilidad de los entornos de aprendizaje. Se concluye que la inteligencia artificial generativa tiene un alto potencial para optimizar los sistemas de tutoría inteligente y mejorar el proceso de enseñanza y aprendizaje, aunque persisten retos relacionados con la evaluación longitudinal, los sesgos algorítmicos y la validación pedagógica.

Palabras clave: inteligencia artificial generativa; sistemas de tutoría inteligente; educación; grandes modelos de lenguaje; aprendizaje personalizado.

ABSTRACT

The integration of generative artificial intelligence in education has fostered the development of new technologies focused on enhancing personalized learning. In this regard, intelligent tutoring systems have evolved significantly with the introduction of large language models capable of generating educational content, providing adaptive feedback, and engaging in conversational interactions with students. This research aimed to analyze intelligent tutoring systems based on generative artificial intelligence through a systematic literature review. The methodology was developed following the PRISMA protocol, which enabled rigorous organization of the search, selection, and analysis stages. Searches were conducted in the Web of Science, IEEE Xplore, and SciELO databases using search equations with the Boolean operators AND and OR. During the identification stage, 1 485 records from the period 2020–2026 were collected and refined by removing duplicates and applying inclusion and exclusion criteria. Nine scientific studies were selected after reviewing titles, abstracts, and full texts according to the established methodological criteria. The results show that intelligent tutoring systems using generative artificial intelligence enable personalized learning, automatic creation of educational materials, personalized feedback, and improved student–system interaction. The incorporation of large language models and generative architectures such as RAG and multi-agent systems enhances the interaction and adaptability of learning environments. It is concluded that generative artificial intelligence has great potential for optimizing intelligent tutoring systems and improving the teaching and learning process, although challenges related to longitudinal evaluation, algorithmic bias, and pedagogical validation remain.

Keywords: generative artificial intelligence; intelligent tutoring systems; education; large language models; personalized learning.

INTRODUCCIÓN

La acelerada transformación digital experimentada por la educación durante la última década ha redefinido profundamente las dinámicas de enseñanza y aprendizaje a escala global. La inteligencia artificial ha orientado sus aplicaciones educativas hacia la mejora de la calidad, la equidad y la personalización de los procesos de aprendizaje. Las organizaciones internacionales han señalado que, cuando se aplican de forma responsable, las tecnologías inteligentes tienen el potencial de mejorar la inclusión y el acceso a los recursos educativos, independientemente del nivel educativo. La UNESCO (2023) reconoce que la integración de la IA en la educación debe enfocarse en el desarrollo de sistemas educativos que ofrezcan

aprendizaje más flexible, adaptable y centrado en el estudiante, por lo que la IA debe ser priorizada en la política educativa.

Conjuntamente, los avances en el aprendizaje profundo (*deep learning*) han aumentado enormemente las capacidades de procesamiento y generación de información en el ámbito digital. En particular, el desarrollo de la inteligencia artificial marca un antes y un después en la creación de sistemas inteligentes aplicables a una amplia variedad de tareas. Bommasani et al. (2021) indican que estos modelos han innovado el manejo de arquitecturas que permiten el procesamiento de contextos complejos y la generación de respuestas coherentes en distintas áreas del conocimiento. De esta forma, los sistemas de IA han evolucionado de ser sistemas tradicionales basados en reglas a sistemas que aprenden e identifican patrones de información para proporcionar respuestas de manera dinámica y en tiempo real.

La IA ha inaugurado una nueva forma de interacción humano-máquina. Estos desarrollos han demostrado un reconocimiento notable de la estructura del lenguaje humano, la capacidad de sostener conversaciones extensas y la habilidad de explicar ideas complejas. OpenAI (2023) menciona que los sistemas generativos son capaces de ajustar sus respuestas en función de la comprensión del usuario, lo que resulta en un gran potencial para su incorporación en procesos de enseñanza y aprendizaje. Por lo tanto, el desarrollo de estas capacidades ha permitido la generación de sistemas de tutoría automatizada que ofrecen retroalimentación contextual, apoyo conceptual y orientación adaptativa durante el desarrollo del proceso educativo.

La rápida expansión de la inteligencia artificial generativa ha suscitado numerosos debates respecto a sus implicaciones éticas y pedagógicas. La literatura internacional ha señalado que la incorporación de estas tecnologías en los procesos educativos requiere la construcción de nuevos marcos normativos que regulen la transparencia, la responsabilización algorítmica y la equidad en el acceso. Holmes et al. (2022) subrayan la necesidad de establecer principios que regulen el uso ético de la IA en la educación. En la misma línea, Zawacki-Richter et al. (2020) plantean que los desarrollos de aplicaciones educativas que incorporen IA deben estar precedidos de una sólida evidencia científica que permita conocer sus efectos en la enseñanza y el aprendizaje.

Dentro de este contexto, los Sistemas de Tutoría Inteligente (ITS, por sus siglas en inglés) se han convertido en una de las aplicaciones más representativas de la inteligencia artificial en educación. Por regla general, estos sistemas se han desarrollado en torno a tres bloques principales: el modelo de dominio, el modelo del estudiante y el modelo pedagógico del tutor. Gracias a estas estructuras, los ITS han podido diagnosticar el conocimiento del estudiante y ofrecer ayuda en el aprendizaje en función de sus necesidades.

Nye (2021) mostró que, a pesar de que estos sistemas presentaban limitaciones en la formulación de explicaciones abiertas y la adaptación semántica profunda, evidenciaban mejoras significativas en el aprendizaje en comparación con los sistemas de enseñanza asistida por computadora de la época. El uso de técnicas de aprendizaje profundo aportó a los sistemas de tutoría una mayor capacidad de adaptación. VanLehn (2021) explica que con la evolución hacia sistemas contemporáneos basados en datos se lograron mejores precisiones para el

diagnóstico del desempeño estudiantil y una mayor capacidad de personalización del aprendizaje. Sin embargo, los modelos basados en datos no lograron resolver la flexibilidad del discurso ni la generación de explicaciones contextualmente pertinentes, lo que demuestra la importancia de contar con tecnologías que posibiliten la integración de interacciones pedagógicas en sus múltiples complejidades.

La personalización en el aprendizaje es uno de los temas más relevantes en el campo educativo. La confluencia de los análisis de aprendizaje y las redes neuronales ha demostrado poder anticipar acciones pedagógicas de forma proactiva. Khosravi et al. (2022) ofrecen evidencia que apoya el uso de estos enfoques para la provisión de retroalimentación diferenciada y la creación de rutas de aprendizaje que se adaptan dinámicamente. Luckin et al. (2022) sostienen que la inteligencia artificial tiene el potencial de actuar como amplificador cognitivo que ayude en la autorregulación del aprendizaje y promueva la personalización de las experiencias educativas.

Los últimos años han visto la convergencia entre los sistemas de tutoría tradicionales y los modelos generativos, dando lugar a una nueva generación de entornos educativos inteligentes. Los avances en los modelos de lenguaje han permitido incorporar en los sistemas de tutoría habilidades conversacionales, generación automática de contenido y simulaciones educativas interactivas. Kasneci et al. (2023) afirman que los modelos generativos pueden producir explicaciones personalizadas, ejemplos contextualizados y diálogos pedagógicos adaptativos que mejoran la comprensión conceptual del aprendiz.

Wang et al. (2024) presentan evidencia empírica positiva respecto al uso de sistemas generativos como tutores digitales, documentando mejoras en la comprensión conceptual mediante la aplicación de retroalimentación explicativa cognitiva y adaptativa, así como el fortalecimiento de los procesos de aprendizaje profundo con el uso de inteligencia artificial generativa. No obstante, los autores subrayan la importancia de nuevos estudios con enfoque longitudinal que permitan valorar el impacto sostenible de estos sistemas en la autodeterminación académica y el desarrollo de habilidades cognitivas de orden superior.

Ante la situación expuesta, los sistemas de tutoría inteligente cobran mayor relevancia en el contexto de la educación digital, al intentar replicar la labor de un tutor mediante un algoritmo que modela el conocimiento, evalúa el desempeño de los estudiantes y brinda retroalimentación personalizada. Rodríguez Chávez (2021) señala que los ITS son capaces de realizar un diagnóstico del aprendizaje del estudiante y ofrecer asistencia académica dentro y fuera del entorno de clase, lo cual favorece la adquisición de habilidades de resolución de problemas y el aprendizaje autónomo en la educación superior.

La expansión de la inteligencia artificial generativa ha incrementado sustancialmente las funcionalidades de los sistemas de tutoría inteligente. Estos sistemas ahora pueden crear contenido educativo dinámico, responder a los estudiantes en forma de diálogo, formular preguntas adaptativas de forma interactiva y en tiempo real, y proporcionar retroalimentación inmediata. Esto ha permitido generar entornos de aprendizaje personalizado de mayor calidad. Maity y Deroy (2024) destacan la forma en que la interacción de los estudiantes con los sistemas de tutoría ha mejorado gracias a la IA generativa.

La investigación en sistemas de tutoría ha incorporado la inteligencia artificial generativa y los agentes de IA con el fin de expandir la funcionalidad de estos sistemas. As'ad (2024) desarrolló una arquitectura pionera asociada a la IA generativa y agentes de IA capaces de proporcionar retroalimentación adaptable en tiempo real de forma interactiva. Esto valida y mejora la pedagogía del contenido generado dinámicamente por la IA, y amplía la capacidad de los sistemas de tutoría para ser utilizados en diversos contextos educativos.

A pesar de estos avances, la literatura científica también ha identificado desafíos relevantes. Maurya y Kochmar (2025) señalan la posibilidad de que los sistemas tutoriales generativos proporcionen tutoría altamente personalizada similar a la mediación humana, pero advierten sobre el potencial sesgo algorítmico y el desafío de cuantificar el impacto educativo de estas tecnologías, junto con la ausencia de marcos estandarizados para la evaluación pedagógica. La convergencia de la inteligencia artificial generativa y los sistemas de tutoría inteligente crea un nuevo paradigma educativo caracterizado por la personalización extrema del aprendizaje, la elaboración contextual de contenidos y la mediación cognitiva. Sin embargo, la consolidación de este campo emergente demanda un estudio a fondo de la evidencia empírica existente. Considerando este panorama, se justifica la realización de una revisión sistemática siguiendo el protocolo PRISMA para analizar el estado del arte en los sistemas de tutoría inteligente basados en inteligencia artificial generativa y, en consecuencia, orientar la prospectiva de investigación y las aplicaciones tecnológicas en el campo de la educación.

MATERIALES Y MÉTODOS

La investigación se llevó a cabo a través de una revisión sistemática de la literatura, desarrollada de acuerdo con las directrices metodológicas del protocolo PRISMA (Elementos Preferidos para Informes de Revisiones Sistemáticas y Metaanálisis), que permitió organizar, identificar, seleccionar y analizar estudios científicos de manera transparente y replicable. El objetivo fue identificar los estudios relevantes relacionados con sistemas de tutoría inteligente basados en inteligencia artificial generativa y analizar sus características metodológicas, aplicaciones educativas, variables estudiadas y hallazgos principales en la literatura científica reciente.

El proceso de búsqueda se realizó en bases de datos científicas de alto impacto: Web of Science, IEEE Xplore y SciELO, seleccionadas por su capacidad de indexación y rigor académico. La búsqueda empleó ecuaciones de búsqueda estructuradas con los operadores booleanos AND y OR, que permiten la combinación de diferentes términos clave relacionados con el tema del estudio. La Tabla 1 presenta las ecuaciones de búsqueda y los resultados obtenidos en cada base de datos.

Tabla 1

Estrategia de búsqueda por base de datos

Base de datos	Ecuaciones de búsqueda aplicadas	Resultados obtenidos
Web of Science	Inteligencia artificial generativa	265
	Inteligencia artificial generativa AND tutoría	5
	Intelligent tutoring systems AND generative artificial intelligence	1 106
	Tutoría académica OR inteligencia artificial generativa	3
IEEE Xplore	Intelligent tutoring systems AND generative artificial intelligence	16
	Generative artificial intelligence AND education	2
SciELO	Inteligencia artificial generativa	82
	Tutoría académica OR inteligencia artificial	6

Nota. WoS = Web of Science. Las ecuaciones de búsqueda se aplicaron en los campos de título, resumen y palabras clave.

Durante la fase de identificación, los filtros temáticos aplicados a la inteligencia artificial generativa, los sistemas de tutoría inteligente y la educación produjeron un total inicial de 1 485 registros científicos de las tres bases de datos analizadas: 1 379 en Web of Science, 18 en IEEE Xplore y 88 en SciELO. Estos resultados formaron el corpus inicial para el posterior proceso de revisión y selección de la literatura. Estos resultados evidenciaron una mayor concentración de investigaciones en bases de datos internacionales de alto impacto, particularmente en Web of Science, lo cual confirma el creciente interés académico por el uso de inteligencia artificial generativa en sistemas de tutoría inteligente.

El protocolo PRISMA se aplicó en cuatro fases secuenciales, tal como se describe en la Tabla 2.

Tabla 2

Fases del protocolo PRISMA aplicado en la revisión sistemática

Fase PRISMA	Procedimiento aplicado	Resultado	Observaciones metodológicas
Identificación	Búsqueda en WoS, IEEE Xplore y SciELO	Registros iniciales (n=1 485)	Uso de operadores booleanos AND y OR
Identificación	Eliminación de duplicados	Registros depurados (n=1 125)	Uso de gestor bibliográfico Mendeley
Cribado	Revisión de títulos y resúmenes	Registros excluidos (n=1 029); registros elegibles (n=96)	Aplicación de criterios de inclusión/exclusión

Fase PRISMA	Procedimiento aplicado	Resultado	Observaciones metodológicas
Elegibilidad	Lectura a texto completo	Artículos excluidos con justificación (n=12)	Motivos documentados por los revisores
Inclusión	Estudios finales incluidos	Estudios analizados (n=9)	Cumplen criterios metodológicos establecidos

Nota. WoS = Web of Science. La eliminación de duplicados se realizó mediante el gestor bibliográfico Mendeley.

Tras la depuración de registros, se inició la fase de cribado mediante la revisión de títulos y resúmenes para determinar la relevancia de los estudios respecto a los objetivos de la investigación. En esta etapa, 1 029 estudios fueron excluidos por no tratar directamente sobre inteligencia artificial generativa en sistemas de tutoría o por encontrarse fuera del ámbito educativo, quedando 96 estudios para la fase de elegibilidad.

El proceso de elegibilidad consistió en la lectura detenida de cada uno de los artículos seleccionados, evaluando el grado de cumplimiento de los criterios de inclusión previamente establecidos. Dentro de estos criterios, se consideró que los estudios debían mencionar de manera explícita la inclusión de alguna forma de inteligencia artificial generativa o modelos de lenguaje en sistemas de tutoría inteligente, aportar información metodológica suficiente y reportar variables educativas tales como el nivel educativo, tipo de IA, diseño de la investigación, tamaño de la muestra, variables analizadas, hallazgos y limitaciones. En esta etapa se excluyeron 12 artículos por no ajustarse a los criterios metodológicos o por no aportar suficiente información para su consideración. Los criterios de inclusión y exclusión se detallan en la Tabla 3.

Tabla 3

Criterios de inclusión y exclusión aplicados en la revisión sistemática

Criterios de inclusión	Criterios de exclusión
Publicaciones entre 2020 y 2026	Publicaciones anteriores a 2020
Artículos revisados por pares	Opiniones o editoriales sin evidencia empírica
Estudios sobre ITS con IA generativa	ITS tradicionales sin componente de IA generativa
Contexto educativo formal	Aplicaciones fuera del ámbito educativo

Nota. ITS = sistemas de tutoría inteligente (Intelligent Tutoring Systems).

En la etapa de inclusión, nueve estudios científicos cumplieron la totalidad de los criterios establecidos y fueron objeto de análisis más detallado mediante la matriz de extracción

de datos. Estos artículos permitieron analizar variables como el nivel educativo de aplicación, los tipos de inteligencia artificial generativa, los diseños metodológicos, las muestras analizadas, los hallazgos principales y las limitaciones reportadas, proporcionando un panorama actualizado sobre la investigación en sistemas de tutoría inteligente basados en inteligencia artificial generativa.

ANÁLISIS Y DISCUSIÓN DE RESULTADOS

A continuación, se detallan los resultados de la revisión sistemática considerando las principales variables extraídas de los nueve estudios seleccionados: nivel educativo de aplicación, tipo de inteligencia artificial generativa, diseño metodológico, muestra analizada, variables educativas, principales hallazgos y limitaciones reportadas por los autores. El análisis de estas variables facilita la comprensión de las tendencias de la investigación y el alcance, impacto y potencial de los sistemas de tutoría inteligente basados en inteligencia artificial generativa en los procesos de enseñanza y aprendizaje en los espacios educativos digitales. La Tabla 4 presenta la matriz de extracción de datos con los nueve estudios incluidos en la revisión.

Tabla 4

Matriz de extracción de datos de los estudios incluidos (n=9)

Autor (año)	Nivel educativo	Tipo de IA generativa	Diseño metodológico	Muestra	Variables analizadas	Principales hallazgos	Limitaciones
Menacho Ángeles, M. R., Pizarro Arancibia, L. M., Osorio Menacho, J. A., & León Pizarro, B. L. (2024)	Educación superior	IA generativa: chatbots, ChatGPT, analítica de aprendizaje, tutor virtual	Cuantitativa descriptiva transversal con revisión documental	200 estudiantes universitarios	Aprendizaje autónomo, motivación, satisfacción, implementación de IA	La IA potencia el aprendizaje autónomo al optimizar la búsqueda de información, asistencia virtual, tutoría inteligente y evaluación automatizada. Se requiere uso con criterios éticos y pedagógicos.	Muestreo no probabilístico por conveniencia; resultados limitados al contexto estudiado; se necesita más investigación sobre aplicaciones educativas de la IA.
Almetnawy, H., Orabi, A., Alneyadi, A. R., Ahmed, T., & Lakas, A. (2025)	Educación superior / aprendizaje digital	GPT y LLM con sistemas multiagente (MAS) y prompt engineering	Diseño y desarrollo de sistema ITS con validación mediante pruebas experimentales con estudiantes simulados	Estudiantes simulados por IA con diferentes grados de competencia	Personalización, adaptabilidad, interacción alumno-IA, detección de fatiga, motivación por gamificación, retroalimentación adaptativa	La IA generativa en ITS permite tutorías personalizadas, mejora la adaptabilidad a distintos perfiles de estudiante e incrementa la motivación mediante gamificación, simulaciones interactivas y retroalimentación adaptativa.	Validación con estudiantes simulados por IA, no con estudiantes reales; se requiere investigación de campo para evaluar el impacto en contextos reales de enseñanza-aprendizaje.

Autor (año)	Nivel educativo	Tipo de IA generativa	Diseño metodológico	Muestra	Variables analizadas	Principales hallazgos	Limitaciones
Calo, T., & MacLellan, C. J. (2024)	Educación superior y entornos de aprendizaje digital	Grandes Modelos de Lenguaje (LLM) con prompt engineering para generación de interfaces de tutoría	Investigación de diseño y desarrollo tecnológico con evaluación comparativa a pequeña escala	Evaluación piloto a pequeña escala (participantes no especificados)	Diseño de interfaces de tutores, eficiencia en la creación de ITS, interacción docente-IA, generación automática de contenidos y rutas de aprendizaje	Los docentes pueden usar IA generativa para diseñar interfaces de tutoría inteligente sin conocimientos técnicos especializados, lo que potencialmente incrementa la efectividad y adopción de ITS.	Evaluaciones piloto a escala reducida; falta validación extensa con docentes y estudiantes en contextos reales de enseñanza-aprendizaje.
Frank, L., Herth, F., Stuwe, P., Klaiber, M., Gerschner, F., & Theissler, A. (2024)	Educación superior (programación en cursos STEM)	LLM: GPT-3.5 (ChatGPT), GPT-4 (Bing), PaLM 2 (Google Bard)	Estudio experimental cuantitativo comparando modelos de lenguaje en un ITS para programación	Evaluación experimental en 4 tipos de entornos de problemas de programación en R (sin reporte directo del tamaño de muestra estudiantil)	Desempeño de los LLM en tutoría de programación, calidad de respuestas, refuerzo del aprendizaje, comparación de modelos	Los ITS con LLM pueden asistir el aprendizaje de programación en R. El modelo GPT-3.5 explicó y resolvió mejor las tareas de utilidad para el estudiante.	Evaluación centrada en el rendimiento del modelo, no en el aprendizaje real de los estudiantes; ausencia de validación educativa longitudinal; dependencia de la calidad del modelo y de los prompts.
Michuy Rodas, F. D. M., Aquije Ramírez, K., & Chávez Paredes, R. C. (2026)	Educación básica (contexto escolar)	IA aplicada a tutoría educativa: ITS y herramientas de IA educativa	Revisión sistemática según directrices PRISMA	29 estudios publicados entre 2023 y 2025 en Scopus, Web of Science y SciELO	Personalización del aprendizaje, tutoría docente asistida por IA, percepción y capacitación docente, dimensión ética y pedagógica, normatividad institucional	La IA puede ofrecer tutorías personalizadas, optimizar la gestión educativa y generar alertas sobre problemas académicos. Requiere integración de innovación tecnológica, pedagogía crítica y políticas educativas.	Desafíos en formación de maestros en IA, brecha tecnológica, transparencia algorítmica y equidad en el acceso a la tecnología.
Liu, G., Guo, Y., Hu, X., & Zhao, Y. (2024)	Educación superior (pregrado)	GPT-4 (LLM) con marco modular basado en estructuras JSON	Estudio piloto pretest-postest de un solo grupo (1 semana) + cuestionario de experiencia de usuario	n = 30 estudiantes de primer año de universidad (habilidades en inglés; sesiones de 30-40 min)	Vocabulario, gramática y construcción de oraciones; efectividad percibida, compromiso, adaptabilidad, satisfacción y precisión de recomendaciones	El sistema de aprendizaje de inglés (SPL) basado en GPT-4 mostró mejoras en vocabulario, gramática y síntesis, así como altos niveles de compromiso, adaptabilidad y satisfacción. El marco	Tamaño de muestra pequeño; duración corta; ausencia de grupo de control (inferencias causales limitadas); dominio único (inglés); demandas computacionales y necesidad de validación de contenido.

Autor (año)	Nivel educativo	Tipo de IA generativa	Diseño metodológico	Muestra	Variables analizadas	Principales hallazgos	Limitaciones
						modular basado en JSON resultó escalable y reutilizable.	
Liu, Z., Agrawal, P., Singhal, S., Madaan, V., Kumar, M., & Verma, P. (2025)	Educación superior / aprendizaje digital personalizado	LLM (GPT-3.5) + RAG (Generación Aumentada por Recuperación) + ingeniería de prompts	Diseño y evaluación de un sistema de tutoría inteligente con IA generativa; evaluación del desempeño del modelo	Evaluación del sistema mediante análisis de las respuestas generadas y revisión por expertos (sin informe extenso de muestra estudiantil)	Precisión, integridad, claridad, coherencia, relevancia y alineación del nivel de dificultad	El sistema LPITutor genera respuestas adaptativas a cada estudiante y logra un equilibrio en precisión, claridad y adecuación pedagógica. Se espera que los LLM continúen mejorando en el aprendizaje adaptativo.	Riesgo de respuestas incorrectas del LLM; desafíos éticos (sesgo, privacidad de datos); necesidad de validación pedagógica y mejoras en la adaptación del sistema.
Levonian, Z., Henkel, O., Li, C., & Postle, M.-E. (2025)	Educación básica secundaria (K-12)	LLM GPT-3.5 integrado en ITS con RAG (Generación Aumentada por Recuperación)	Estudio mixto: prueba de usabilidad en aula y despliegue en campo; evaluación empírica en dos momentos	109 estudiantes en pruebas en aula; más de 8 168 usuarios y 126 000 mensajes analizados en el despliegue masivo	Seguridad del sistema, calidad de respuestas del LLM, interacción estudiante-chatbot, moderación de contenido, fundamentación de respuestas	La integración de LLM con ITS mejora las interacciones personalizadas y apoya la comprensión conceptual en matemáticas. RAG junto con sistemas de moderación mejoran la calidad y precisión; solo el 0,31% de mensajes contenía contenido problemático.	Riesgo de alucinaciones del modelo; posibles sesgos culturales del LLM; necesidad de mejorar la moderación automática; limitaciones en la generalización a otros contextos educativos; falta de evaluación directa del impacto en el aprendizaje a largo plazo.
Zhang, L., Lin, J., Borchers, C., Cao, M., & Hu, X. (2024)	Educación superior y aprendizaje digital en ITS	GAN (Generative Adversarial Networks) y GPT (Generative Pre-trained Transformer)	Estudio experimental computacional con desarrollo del marco algorítmico 3DG para generación y densificación de datos de aprendizaje	Datos de rendimiento estudiantil del sistema AutoTutor (Centro para el Estudio de la Alfabetización de Adultos - CSAL)	Desempeño en cuestionarios, número de intentos, modelado del estudiante, datos sintéticos para el aprendizaje	El marco 3DG mejora la simulación y generación de datos en ITS cuando los datos son escasos. Para la generación de datos educativos, GAN exhibió mayor fiabilidad que GPT-4.	Dependencia de conjuntos de datos específicos del sistema AutoTutor; falta de validación en múltiples contextos educativos; necesidad de evaluar el impacto pedagógico real.

Nota. ITS = intelligent tutoring systems; LLM = large language models; RAG = Retrieval-Augmented Generation (Generación Aumentada por Recuperación); GAN = Generative Adversarial Networks; MAS = multi-agent systems.

Según la evidencia sintetizada bajo PRISMA, los Sistemas de Tutoría Inteligente (ITS) basados en inteligencia artificial generativa se han establecido como un desarrollo educativo emergente con dos vías dominantes: (a) tutoría conversacional a través de los LLMs (GPT-3.5, GPT-4 y variantes diseñadas con prompts), centrada en la interactividad, retroalimentación y personalización; y (b) IA generativa como soporte del ecosistema ITS (generación de datos sintéticos, densificación de registros, análisis y automatización de componentes), con un énfasis en abordar limitaciones estructurales como la escasez de datos y la escalabilidad institucional, tal como señalan los trabajos revisados (Menacho Ángeles et al., 2024; Zhang et al., 2024; Michuy Rodas et al., 2026).

Los resultados evidenciaron que la IA generativa se relaciona con el aprendizaje autónomo, la motivación para aprender y la satisfacción cuando se integra en el entorno virtual como ambiente de gestión de información académica. Esto destaca el papel de la IA generativa como mediador para la tutoría, la asistencia y las evaluaciones automatizadas en la educación superior, cuando se enmarca con un uso ético y parámetros pedagógicos definidos, como reporta la evidencia cuantitativa transversal con estudiantes universitarios (Menacho Ángeles et al., 2024). La revisión sistemática PRISMA del período 2023–2025 también incorporó las percepciones institucionales y docentes que sitúan a la IA en los roles de personalización, advertencia temprana y acompañamiento, dependientes de la formación docente y de una gobernanza educativa explícita (Michuy Rodas et al., 2026).

En cuanto a la revisión de la literatura, los LLMs se sitúan a la vanguardia de los ITS contemporáneos por su capacidad para generar explicaciones, mantener diálogos y adaptar su discurso. Sin embargo, su rendimiento se ha medido con frecuencia en términos de métricas de calidad de respuesta en lugar del aprendizaje real de los estudiantes.

Los estudios comparativos se centraron en el rendimiento del modelo al ejecutar tareas de programación y resolución de problemas, e informaron una utilidad práctica para guiar los procesos de aprendizaje, con ventajas relativas dependiendo del modelo y el diseño del prompt, como se demostró en el trabajo experimental sobre programación (Frank et al., 2024) y en la evaluación del sistema LPITutor (Liu Z. et al., 2025). La evidencia de medidas directas de aprendizaje con estudiantes reales fue más limitada, aunque proporcionó señales consistentes. Un estudio piloto pretest-postest de una semana con estudiantes universitarios reportó mejoras en habilidades de inglés (vocabulario, gramática y construcción de oraciones), así como altas calificaciones de compromiso, adaptabilidad y satisfacción de uso, asociadas a un marco modular y estructurado; sin embargo, el diseño de grupo único limita la atribución causal (Liu, Guo, Hu & Zhao, 2024).

De manera transversal, las arquitecturas de apoyo resultan fundamentales para mejorar la confiabilidad y el control pedagógico. Los sistemas con RAG (Generación Aumentada por Recuperación) integran la fundamentación documental y reducen los riesgos de respuestas inapropiadas, obteniendo resultados robustos a gran escala en matemáticas: un despliegue masivo que analizó más de 8 168 usuarios y 126 000 mensajes reportó baja incidencia de problemas cuando se establecieron límites y moderación, junto con mejoras en la calidad atribuidas al anclaje por recuperación (Levonian et al., 2025). Simultáneamente, LPITutor

reportó un equilibrio entre exactitud, claridad y adecuación al nivel de dificultad, aunque los riesgos de error del modelo y los problemas éticos relacionados con la datificación y los sesgos persisten (Liu Z. et al., 2025).

Otra frontera crítica se ubica en la calidad de los datos educativos. En los ITS, la ausencia de un registro de desempeño suficiente limita la modelización del estudiante y la evaluación de su conocimiento; el marco 3DG abordó este cuello de botella mediante la tensorización, la factorización y el aumento con modelos generativos, logrando simulaciones contextualizadas en el sistema AutoTutor. El análisis comparativo informó una mayor fiabilidad de las GAN sobre GPT-4 para la generación de datos educativos en este contexto, reforzando el argumento de que la IA generativa no solo "tutoriza", sino que también fortalece la dimensión analítica de los ITS (Zhang et al., 2024).

Se observó también una tendencia clara hacia la democratización del diseño de los ITS. El uso de LLMs para permitir a los educadores diseñar interfaces de tutoría y componentes sin requerir programación avanzada facilita la creación rápida de tutores personalizados mediante generación automática de diseño y componentes interactivos, con evaluaciones preliminares de usuarios a pequeña escala (Calo & MacLellan, 2024). Este resultado está directamente relacionado con la necesidad institucional de minimizar las barreras técnicas y mejorar la capacidad de innovación pedagógica, aunque aún no se ha abordado la validación extensa con docentes y estudiantes en condiciones reales.

La variedad de estudios reveló una fortaleza adicional derivada del uso de un enfoque metodológico mixto: estudios descriptivos, encuestas, desarrollos computacionales, estudios experimentales orientados al rendimiento, estudios piloto pretest-postest y revisiones sistemáticas PRISMA coexistieron en el corpus analizado. Esta diversidad fortaleció el mapeo del campo, aunque la comparabilidad directa de los hallazgos fue más débil debido a la variación en las unidades de análisis (estudiantes reales, estudiantes simulados por IA, revisores expertos), la duración (sesiones cortas vs. implementaciones masivas) y el rango de variables estudiadas (aprendizaje, usabilidad, seguridad, calidad textual, precisión, motivación), tal como demostraron los estudios de Almetnawy et al. (2025), Frank et al. (2024), Liu et al. (2024) y Levonian et al. (2025).

El análisis de limitaciones reveló cuatro brechas consistentes a lo largo de los estudios. Primero, ausencia de inferencia causal robusta por falta de grupos de control, tamaños de muestra reducidos y duraciones cortas (Liu et al., 2024). Segundo, elevado número de evaluaciones centradas en la "calidad de respuesta" del modelo sin medir el aprendizaje longitudinal (Frank et al., 2024; Liu Z. et al., 2025). Tercero, generalización pedagógica limitada, dado que varios desarrollos se validaron con estudiantes simulados o muestras sin suficiente caracterización (Almetnawy et al., 2025; Calo & MacLellan, 2024). Cuarto, persistencia de desafíos técnico-éticos asociados con el sesgo, la privacidad y la transparencia algorítmica (Michuy Rodas et al., 2026; Liu Z. et al., 2025).

La discusión integrada permite concluir que los ITS con IA generativa muestran un potencial real para la personalización, escalabilidad e interacción pedagógica, especialmente cuando se combinan con RAG, moderación y diseño centrado en la seguridad (Levonian et al.,

2025). La evidencia complementaria confirma que el avance del campo depende tanto del tutor conversacional como de la calidad de los datos y del modelado del estudiante, donde enfoques como 3DG refuerzan el soporte analítico de los ITS (Zhang et al., 2024). Las revisiones PRISMA y los estudios empíricos alinean una agenda clara: evaluación educativa longitudinal con diseño de grupo comparativo, formación docente en IA, políticas institucionales sobre gobernanza y marcos integrativos que combinen ética, pedagogía y arquitectura técnica en el diseño (Menacho Ángeles et al., 2024; Michuy Rodas et al., 2026).

CONCLUSIONES

- La revisión sistemática evidencia que los sistemas de tutoría inteligente integrados con inteligencia artificial generativa constituyen una de las tecnologías más relevantes en la educación contemporánea, debido a que facilitan aprendizajes individualizados, la generación automatizada de contenidos, la retroalimentación en tiempo real y la mejora de la interacción de los estudiantes con los entornos de aprendizaje digitales.
- Los estudios revisados confirman que la mejora de los Modelos de Lenguaje a Gran Escala (LLM), las arquitecturas de inteligencia artificial generativa y los sistemas mixtos como RAG (Generación Aumentada por Recuperación) o sistemas multiagente potencian de manera significativa los sistemas de tutoría inteligente en el aprendizaje adaptativo, la tutorización automática y la comprensión de conceptos en diversas áreas del conocimiento.
- A pesar de los avances tecnológicos documentados en la literatura, persisten limitaciones y retos relevantes, tales como la carencia de estudios de impacto longitudinal en el aprendizaje, el uso de muestras pequeñas o simuladas y la ausencia de marcos de evaluación pedagógica que permitan medir la efectividad de estos sistemas en el aprendizaje a lo largo del tiempo.
- La mejora en la implementación de sistemas de tutoría inteligente con inteligencia artificial generativa debe basarse, de forma equilibrada, en la innovación tecnológica, la teoría pedagógica, la ética y la investigación empírica sobre los entornos educativos reales, para consolidar el uso de estos sistemas en la mejora y personalización de la experiencia de aprendizaje en la educación digital.

REFERENCIAS BIBLIOGRÁFICAS

- Almetnawy, H., Orabi, A., Alneyadi, A. R., Ahmed, T., & Lakas, A. (2025). Generative AI-driven intelligent tutoring systems with multi-agent architecture and adaptive prompt engineering. En *Proceedings of the IEEE Global Engineering Education Conference (EDUCON 2025)*. IEEE.
- As'ad, M. (2024). Generative artificial intelligence in intelligent tutoring systems: A proof-of-concept for adaptive learning and root cause analysis in healthcare education. *Cureus*. <https://doi.org/10.7759/cureus.69710>
- Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., ... Liang, P. (2021). On the opportunities and risks of foundation models. *arXiv preprint*. <https://arxiv.org/abs/2108.07258>
- Calo, T., & MacLellan, C. J. (2024). Generating intelligent tutoring system interfaces using large language models. *arXiv*. <https://doi.org/10.48550/arXiv.2405.14713>

- Frank, L., Herth, F., Stuwe, P., Klaiber, M., Gerschner, F., & Theissler, A. (2024). Evaluating large language models for intelligent tutoring systems in programming education. En Proceedings of the IEEE Global Engineering Education Conference (EDUCON 2024). IEEE. <https://doi.org/10.1109/EDUCON60312.2024.10578933>
- Holmes, W., Bialik, M., & Fadel, C. (2022). Artificial intelligence in education: Promises and implications for teaching and learning. Center for Curriculum Redesign.
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Khosravi, H., Kitto, K., & Gašević, D. (2022). The future of learning analytics: Artificial intelligence and learning design. *British Journal of Educational Technology*, 53(6), 1467–1475. <https://doi.org/10.1111/bjet.13224>
- Levonian, Z., Henkel, O., Li, C., & Postle, M.-E. (2025). Deploying generative AI chat tutors in intelligent tutoring systems: Safety, interaction quality and large-scale usage analysis [Preprint]. Zenodo. <https://doi.org/10.5281/zenodo.8284412>
- Liu, G., Guo, Y., Hu, X., & Zhao, Y. (2024). Personalized language learning with GPT-4: A modular intelligent tutoring system for English language development. *Electronics*, 13(24), 4876. <https://doi.org/10.3390/electronics13244876>
- Liu, Z., Agrawal, P., Singhal, S., Madaan, V., Kumar, M., & Verma, P. (2025). LPITutor: A personalized intelligent tutoring system based on large language models using retrieval-augmented generation and prompt engineering. *PeerJ Computer Science*, 11, e2991. <https://doi.org/10.7717/peerj-cs.2991>
- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. (2022). Intelligence unleashed: An argument for AI in education. Pearson Education.
- Maity, S., & Deroy, A. (2024). Generative artificial intelligence for intelligent tutoring systems: Opportunities and challenges. arXiv. <https://doi.org/10.48550/arXiv.2410.10650>
- Maurya, K. K., & Kochmar, E. (2025). Pedagogy-driven evaluation of generative AI-powered intelligent tutoring systems. arXiv. <https://doi.org/10.48550/arXiv.2510.22581>
- Menacho Ángeles, M. R., Pizarro Arancibia, L. M., Osorio Menacho, J. A., & León Pizarro, B. L. (2024). Inteligencia artificial generativa y aprendizaje autónomo en educación superior [Preprint]. Zenodo. <https://doi.org/10.5281/zenodo.10693945>
- Michuy Rodas, F. D. M., Aquije Ramírez, K., & Chávez Paredes, R. C. (2026). Inteligencia artificial y tutoría educativa: Revisión sistemática bajo el protocolo PRISMA [Artículo en prensa/preprint]. Zenodo. <https://doi.org/10.5281/zenodo.17118338>
- Nye, B. D. (2021). Intelligent tutoring systems by and for the real world: A review of recent advances. *International Journal of Artificial Intelligence in Education*, 31(4), 678–704. <https://doi.org/10.1007/s40593-020-00236-1>
- OpenAI. (2023). GPT-4 technical report. arXiv preprint. <https://arxiv.org/abs/2303.08774>
- Rodríguez Chávez, M. H. (2021). Sistemas de tutoría inteligente y su aplicación en la educación superior. *Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, 11(22). <https://doi.org/10.23913/ride.v11i22.848>

- UNESCO. (2023). Guidance for generative AI in education and research. United Nations Educational, Scientific and Cultural Organization. <https://unesdoc.unesco.org/ark:/48223/pf0000386693>
- VanLehn, K. (2021). Two principles for improving AI in education. *International Journal of Artificial Intelligence in Education*, 31(2), 238–248. <https://doi.org/10.1007/s40593-020-00221-8>
- Wang, Y., Li, X., Chen, G., & Zhang, J. (2024). Generative AI-powered tutoring systems and student conceptual understanding: An experimental study. *Computers & Education: Artificial Intelligence*, 5, 100162. <https://doi.org/10.1016/j.caeai.2024.100162>
- Zawacki-Richter, O., Bond, M., Marin, V. I., & Gouverneur, F. (2020). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? *International Journal of Educational Technology in Higher Education*, 17, 39. <https://doi.org/10.1186/s41239-020-00212-8>
- Zhang, L., Lin, J., Borchers, C., Cao, M., & Hu, X. (2024). 3DG: A framework for using generative AI to handle sparse student performance data from intelligent tutoring systems. *arXiv*. <https://doi.org/10.48550/arXiv.2402.01746>



Casa Editora del Polo (CASEDELPO), hace constar que:

El artículo científico:

**Sistemas de tutoría inteligente basados en inteligencia artificial generativa: una
revisión sistemática según directrices PRISMA**

De autoría:

Edison Guillermo Vergara Aguilera, Luis Eduardo Solis-Granda

Habiéndose procedido a su revisión y analizados los criterios de evaluación realizados por lectores pares expertos (externos) vinculados al área de experticia del artículo presentado, ajustándose el mismo a las normas que comprenden el proceso editorial, se da por aceptado la publicación en el **Vol. 11, No 2, 2026**, de la revista Polo del Conocimiento, con ISSN 2550-682X, indexada y registrada en las siguientes bases de datos y repositorios: **Latindex Catálogo v2.0, MIAR, Google Académico, ROAD, Dialnet, ERIHPLUS.**

Y para que así conste, firmo la presente en la ciudad de Manta, a los **7 de marzo de 2026**.

Dr. Víctor R. Jama Zambrano
DIRECTOR



Casa Editora del Polo (CASEDELPO), hace constar que:

El artículo científico:

**Sistemas de tutoría inteligente basados en inteligencia artificial generativa: una
revisión sistemática según directrices PRISMA**

De autoría:

Edison Guillermo Vergara Aguilera, Luis Eduardo Solis-Granda

Ha sido publicado en el **Vol. 11, No 2, 2026**, de la revista Polo del Conocimiento con
ISSN 2550-682X, indexada y registrada en las siguientes bases de datos y repositorios:

Latindex Catálogo v2.0, MIAR, Google Académico, ROAD, Dialnet, ERIHPLUS.

Disponible en:

URL: <https://polodelconocimiento.com/ojs/index.php/es/article/view/11237>

Y para que así conste, firmo la presente en la ciudad de Manta, a los **7 de marzo de 2026**.

Dr. Víctor R. Jama Zambrano
DIRECTOR

UNEMI

UNIVERSIDAD ESTATAL DE MILAGRO

¡Evolución académica!

@UNEMIEcuador

