UNEMI
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 EDUCACIÓN CON MENCIÓN EN DOCENCIA E INVESTIGACIÓN EN EDUCACIÓN SUPERIOR

TEMA:

Chatbots basados en inteligencia artificial como tutores académicos en la educación superior STEM: una revisión sistemática de la literatura

Autores:

Lenin Fernando Ortiz Pineda

Priscila Elizabeth Mackliff Montero

Director:

Dennis Alfredo Peralta Gamboa

Milagro, 2026

Chatbots basados en inteligencia artificial como tutores académicos en la educación superior STEM: una revisión sistemática de la literatura

Resumen

Los chatbots basados en inteligencia artificial han adquirido creciente relevancia como herramientas de apoyo académico en la educación superior, particularmente en las áreas de ciencia, tecnología, ingeniería y matemáticas. El objetivo de este estudio fue analizar sistemáticamente la producción científica reciente sobre el uso de chatbots de inteligencia artificial como tutores académicos en asignaturas STEM, identificando sus aplicaciones pedagógicas, beneficios, limitaciones, riesgos y vacíos de investigación. Se realizó una revisión sistemática de la literatura siguiendo las orientaciones de PRISMA 2020. La búsqueda se efectuó en Scopus, Web of Science, ERIC, ScienceDirect y Google Scholar. Tras las fases de identificación, cribado y elegibilidad, se incluyeron 7 estudios publicados entre 2023 y 2026. Los resultados permitieron identificar cuatro categorías: chatbots como tutores virtuales, retroalimentación personalizada y apoyo académico, aprendizaje autodirigido y pensamiento crítico, y riesgos, precisión y dependencia de la inteligencia artificial. Los hallazgos muestran que estas herramientas pueden favorecer la tutoría personalizada, la explicación de contenidos, la resolución de problemas y la autonomía estudiantil. Sin embargo, persisten riesgos relacionados con errores conceptuales, dependencia y confianza excesiva. Se concluye que su efectividad depende de una integración pedagógica planificada, mediación docente, verificación crítica de respuestas y lineamientos institucionales para su uso responsable.

Palabras clave: chatbots de inteligencia artificial; tutoría académica; educación superior; educación STEM; inteligencia artificial generativa; aprendizaje autodirigido.

Abstract

Artificial intelligence-based chatbots have gained increasing relevance as academic support tools in higher education, particularly in the fields of science, technology, engineering, and mathematics (STEM). The aim of this study was to systematically analyze recent scientific literature on the use of artificial intelligence chatbots as academic tutors in STEM subjects, identifying their pedagogical applications, benefits, limitations, risks, and research gaps. A systematic literature review was conducted following the PRISMA 2020 guidelines. The search was carried out in Scopus, Web of Science, ERIC, ScienceDirect, and Google Scholar. Following the identification, screening, and eligibility stages, seven studies published between 2023 and 2026 were included. The results revealed four categories: chatbots as virtual tutors; personalized feedback and academic support; self-directed learning and critical thinking; and risks, accuracy, and dependence on artificial intelligence. The findings indicate that these tools can promote personalized tutoring, content explanation, problem-solving, and student autonomy. However, risks related to conceptual errors, dependence, and excessive trust in artificial intelligence remain. It is concluded that their effectiveness depends on planned pedagogical integration, teacher mediation, critical verification of responses, and institutional guidelines for their responsible use.

Keywords: artificial intelligence chatbots; academic tutoring; higher education; STEM education; generative artificial intelligence; self-directed learning.

INTRODUCCIÓN

La inteligencia artificial generativa ha adquirido una presencia creciente en la educación superior, especialmente a partir de la expansión de modelos conversacionales capaces de producir explicaciones, resolver problemas, generar código, ofrecer retroalimentación inmediata y acompañar procesos de aprendizaje autónomo. En este escenario, los chatbots de inteligencia artificial, como ChatGPT y otros sistemas basados en modelos de lenguaje, han comenzado a ser considerados no solo como herramientas de apoyo académico, sino también como posibles tutores virtuales capaces de interactuar con los estudiantes mediante lenguaje natural. Esta transformación resulta particularmente relevante en las asignaturas STEM —ciencia, tecnología, ingeniería y matemáticas—, donde los estudiantes suelen enfrentar contenidos abstractos, alta carga cognitiva, resolución de problemas complejos y necesidad constante de retroalimentación personalizada.

La literatura reciente señala que los chatbots educativos pueden contribuir al aprendizaje mediante asistencia en tareas, explicación de conceptos, apoyo personalizado y desarrollo de habilidades académicas (Labadze et al., 2023). En el contexto STEM, estas posibilidades adquieren mayor importancia porque las asignaturas requieren razonamiento lógico, pensamiento computacional, modelación matemática, análisis experimental y capacidad para aplicar conocimientos en situaciones nuevas. Por ejemplo, Ding et al. (2023) analizaron el uso de ChatGPT como tutor virtual en una clase universitaria de física y evidenciaron que los estudiantes pueden percibirlo como una herramienta útil para resolver dudas; sin embargo, también identificaron problemas relacionados con la confianza excesiva en respuestas que no siempre son correctas. De manera similar, Sun et al. (2024) estudiaron el uso de ChatGPT en programación universitaria y observaron cambios positivos en la percepción de utilidad y

facilidad de uso, aunque los resultados en desempeño no mostraron diferencias estadísticamente significativas frente a modalidades de aprendizaje autodirigido.

El problema central radica en que, aunque los chatbots de IA se están incorporando rápidamente en la educación superior, todavía existe incertidumbre sobre su efectividad real como tutores académicos en asignaturas STEM. Las investigaciones disponibles muestran resultados prometedores, pero también advierten riesgos importantes: dependencia excesiva, errores conceptuales, respuestas imprecisas, debilitamiento del pensamiento crítico, problemas de integridad académica y ausencia de criterios pedagógicos claros para su implementación. Lee et al. (2024), por ejemplo, sostienen que el uso de ChatGPT puede favorecer la autorregulación y el pensamiento de orden superior cuando se integra mediante mecanismos de guía que obligan al estudiante a intentar resolver primero los problemas antes de solicitar ayuda. Esto sugiere que el valor pedagógico del chatbot no depende únicamente de la herramienta, sino del diseño didáctico que regula su uso.

La justificación de esta revisión sistemática se fundamenta en la necesidad de organizar críticamente la evidencia científica reciente sobre los chatbots de IA como tutores académicos en STEM. Si bien existen revisiones amplias sobre chatbots en educación y sobre IA generativa en educación superior, aún se requiere una síntesis específica centrada en asignaturas STEM, debido a las particularidades cognitivas, metodológicas y evaluativas de estas áreas. Además, los estudios recientes en matemáticas, física, química, programación y otras áreas STEM muestran hallazgos heterogéneos respecto a beneficios, limitaciones y condiciones de uso. Pepin et al. (2025) destacan que ChatGPT puede apoyar la comprensión de conceptos matemáticos, la retroalimentación en tiempo real y el aprendizaje autorregulado, pero también advierten sobre inexactitudes, dilemas éticos y riesgo de sobredependencia. En esta misma línea, Yusuf et al. (2024) señalan que la IA generativa plantea tanto oportunidades de reforma educativa como preocupaciones asociadas a la integridad académica y la necesidad de políticas institucionales.

Por tanto, el objetivo de este estudio es analizar sistemáticamente la producción científica reciente sobre el uso de chatbots de inteligencia artificial como tutores académicos en asignaturas STEM de educación superior, identificando sus aplicaciones pedagógicas, beneficios reportados, limitaciones, riesgos y vacíos de investigación. El aporte del estudio consiste en ofrecer una visión organizada y crítica del estado actual del conocimiento, diferenciando entre usos instrumentales de los chatbots y experiencias pedagógicamente estructuradas de tutoría académica. Asimismo, esta revisión puede orientar a docentes, investigadores y responsables institucionales en el diseño de estrategias de integración ética, efectiva y supervisada de la IA en la enseñanza STEM universitaria, evitando enfoques tecnocéntricos y promoviendo modelos donde la inteligencia artificial complemente, pero no sustituya, la mediación docente.

METODOLOGÍA

Diseño del estudio

El presente estudio se desarrolló bajo un enfoque de **revisión sistemática de la literatura**, debido a que busca identificar, seleccionar, analizar y sintetizar la evidencia científica disponible sobre el uso de chatbots basados en inteligencia artificial como tutores académicos en asignaturas STEM de educación superior. Este tipo de diseño permite organizar de manera rigurosa los hallazgos existentes, reconocer tendencias investigativas, comparar resultados y establecer vacíos de conocimiento en torno a un fenómeno educativo emergente. En este caso, la revisión se centró en experiencias relacionadas con herramientas como ChatGPT y otros chatbots educativos aplicados en áreas como matemáticas, física, química, programación, ingeniería y ciencias de la computación.

Protocolo de revisión

Para garantizar transparencia y orden metodológico, el proceso de revisión se orientó mediante la declaración **PRISMA 2020** propuesta por Page et al. (2021), la cual ofrece una guía actualizada para reportar revisiones sistemáticas, considerando las fases de identificación, cribado, elegibilidad e inclusión de estudios. PRISMA permite presentar de forma clara cómo se localizaron los documentos, qué criterios se aplicaron para seleccionarlos y cuáles fueron finalmente incluidos en la revisión.

Fuentes de información y estrategia de búsqueda

La búsqueda de información se realizó en bases de datos académicas reconocidas por su cobertura en educación, tecnología e investigación científica, tales como **Scopus, Web of Science, ERIC, ScienceDirect y Google Scholar**. Se emplearon términos de búsqueda en inglés y español, debido a que la mayor parte de la producción científica sobre inteligencia artificial educativa se encuentra publicada en inglés. Las palabras clave utilizadas fueron: *"AI chatbots"*, *"ChatGPT"*, *"virtual tutor"*, *"academic tutoring"*, *"STEM education"*, *"higher education"*, *"artificial intelligence in education"*, así como sus equivalentes en español: "chatbots de inteligencia artificial", "tutor virtual", "tutoría académica", "educación STEM" y "educación superior". Estos términos se combinaron mediante operadores booleanos como **AND** y **OR**, con el propósito de ampliar y precisar los resultados obtenidos.

Criterios de inclusión y exclusión

Se incluyeron artículos científicos publicados entre **2023 y 2026**, considerando que el uso educativo de chatbots generativos adquirió mayor relevancia académica a partir de la aparición y expansión de herramientas como ChatGPT. También se consideraron estudios empíricos, revisiones sistemáticas y trabajos teóricos que abordaran el uso de chatbots de inteligencia artificial en educación superior, específicamente en asignaturas STEM. Asimismo, se incluyeron investigaciones que analizaran beneficios pedagógicos, retroalimentación inmediata, aprendizaje autónomo, desempeño académico, percepción estudiantil, riesgos, limitaciones o condiciones de implementación docente.

Por el contrario, se excluyeron documentos no académicos, notas de opinión, publicaciones sin revisión científica, estudios centrados exclusivamente en educación básica o media, investigaciones sobre inteligencia artificial sin relación directa con chatbots, y trabajos que no abordaran el contexto universitario o las áreas STEM.

Proceso de selección de estudios

El proceso de selección se organizó en cuatro fases (Ver Figura 1). Primero, en la fase de **identificación**, se recopilaron los documentos encontrados en las bases de datos seleccionadas. Luego, en la fase de **cribado**, se revisaron títulos, resúmenes y palabras clave para descartar estudios duplicados o no relacionados con el objetivo de la revisión. En la fase de **elegibilidad**, se analizaron los textos completos de los artículos potencialmente relevantes, verificando su pertinencia temática, metodológica y contextual. Finalmente, en la fase de **inclusión**, se seleccionaron los estudios que cumplían con los criterios establecidos y que aportaban evidencia significativa sobre el uso de chatbots como tutores académicos en educación superior STEM.

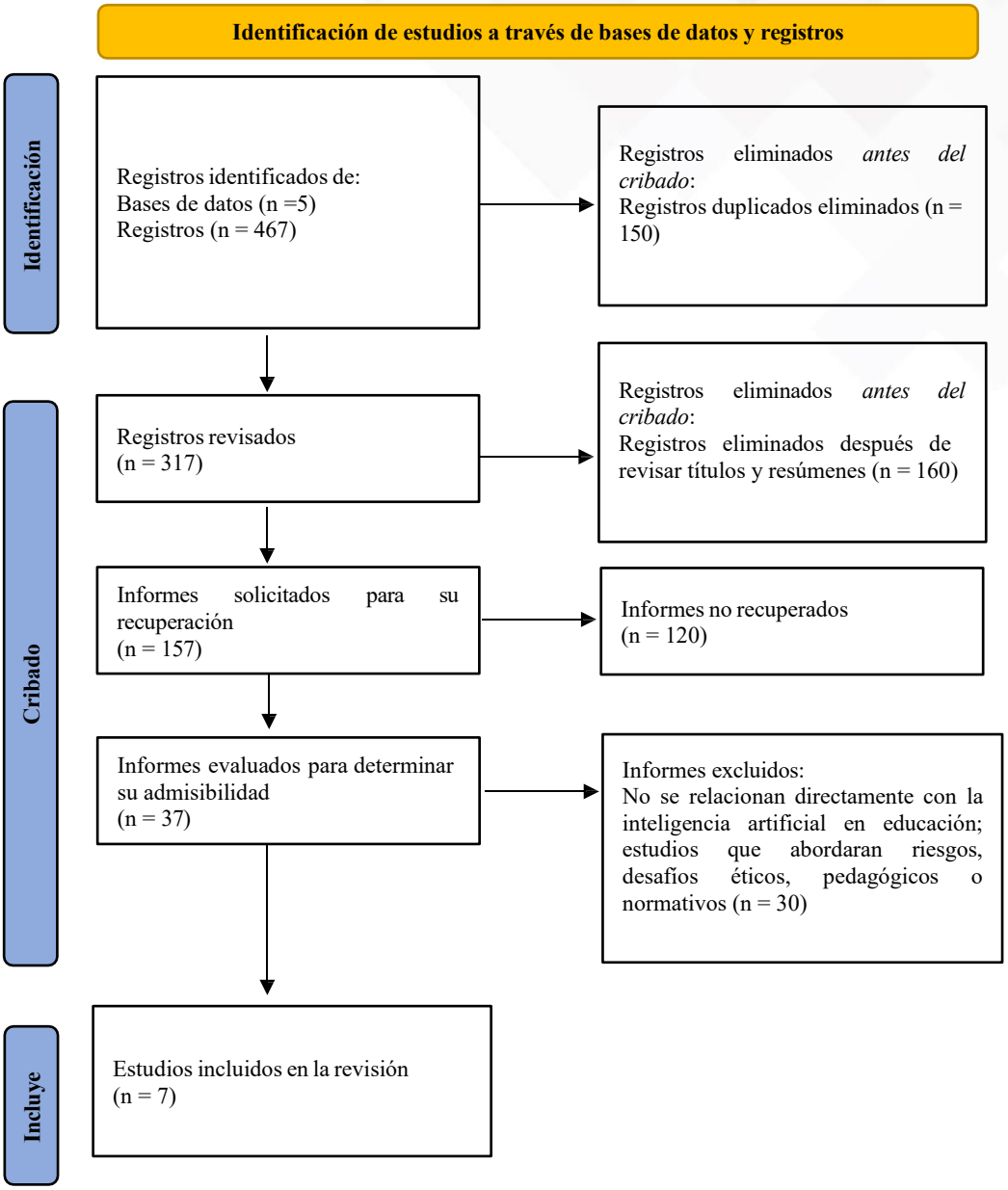


Figura 1. Proceso de selección de estudios

Extracción y análisis de la información

La información de los estudios seleccionados fue organizada en una matriz de análisis que incluyó autor, año de publicación, país o contexto, área STEM abordada, tipo de chatbot utilizado, diseño metodológico, principales hallazgos, beneficios reportados, limitaciones y recomendaciones pedagógicas. Posteriormente, se aplicó una síntesis cualitativa de contenido, agrupando los resultados en categorías temáticas relacionadas con aplicaciones pedagógicas, retroalimentación personalizada, aprendizaje autónomo, desarrollo de habilidades cognitivas, riesgos de dependencia, errores conceptuales, integridad académica y necesidad de mediación docente.

RESULTADOS

Proceso de selección de los estudios

La revisión sistemática permitió identificar investigaciones relacionadas con el uso de chatbots basados en inteligencia artificial, modelos de lenguaje de gran escala, ChatGPT, tutores virtuales, retroalimentación personalizada y apoyo académico en asignaturas STEM de educación superior. El proceso de selección se organizó siguiendo las fases de identificación, cribado, elegibilidad e inclusión, con el propósito de garantizar una revisión ordenada y coherente con el objetivo del estudio.

La búsqueda inicial permitió identificar 467 registros en las bases de datos consultadas. De estos, 150 documentos fueron eliminados por duplicidad, quedando 317 registros para la fase de cribado. Posteriormente, se excluyeron 160 documentos después de revisar títulos, resúmenes y palabras clave, debido a que no respondían directamente al objetivo de la revisión. Luego, se solicitaron 157 informes para su recuperación, de los cuales 120 no fueron recuperados. En la fase de elegibilidad, se evaluaron 37 documentos en texto completo y se excluyeron 30 por no cumplir los criterios de inclusión establecidos. Finalmente, la muestra documental quedó conformada por 7 estudios, los cuales constituyeron la base para el análisis cualitativo de la revisión.

Los estudios incluidos fueron seleccionados porque abordaban de manera directa o complementaria el uso de chatbots, ChatGPT o modelos de lenguaje como recursos de tutoría, apoyo al aprendizaje, retroalimentación o acompañamiento académico en áreas STEM, tales como física, ciencias de la computación, programación, ingeniería química, matemáticas, ciencia de datos e ingeniería.

Caracterización general de los estudios incluidos

Los 7 estudios seleccionados evidencian una producción científica reciente, concentrada entre 2023 y 2026. Este periodo refleja el crecimiento acelerado del interés académico por analizar el papel de los chatbots de inteligencia artificial y los modelos de lenguaje generativo en la educación superior, especialmente después de la expansión de herramientas como ChatGPT. La mayoría de investigaciones se ubica en contextos universitarios y se enfoca en asignaturas STEM, lo que resulta pertinente debido a la complejidad conceptual, procedimental y cognitiva de estas áreas.

Desde el punto de vista metodológico, los estudios presentan enfoques variados. Se identificaron estudios empíricos con estudiantes universitarios, propuestas de diseño e implementación de tutores virtuales, evaluaciones de modelos de lenguaje para tutoría matemática, experiencias con chatbots de recuperación aumentada y análisis del aprendizaje autodirigido apoyado por herramientas GPT. Esta diversidad metodológica permite comprender el fenómeno desde distintas perspectivas: percepción estudiantil, diseño tecnológico, apoyo personalizado, retroalimentación, aprendizaje autónomo y limitaciones pedagógicas.

Tabla 1. Características de los estudios incluidos

Autor/año	Área STEM	Tipo de recurso de IA	Enfoque del estudio	Principales aportes
Ding et al. (2023)	Física	ChatGPT como tutor virtual	Estudio empírico sobre percepción estudiantil	Analiza cómo los estudiantes perciben el uso de ChatGPT como tutor virtual en una clase universitaria de física.
Bassner et al. (2024)	Ciencias de la computación	Iris, tutor virtual basado en IA	Diseño e implementación	Presenta un tutor virtual para apoyar ejercicios de programación y aprendizaje interactivo en ciencias de la computación.
Caccavale et al. (2024)	Ingeniería química	Modelos de lenguaje como tutores virtuales	Estudio aplicado/reflexivo	Examina el papel de los LLM como tutores virtuales en el marco de la Educación 4.0 en ingeniería química.
Gupta et al. (2025)	Matemáticas	Modelos de lenguaje para tutoría matemática	Evaluación de desempeño de LLM	Evalúa la capacidad de los LLM para ofrecer tutoría matemática más allá de entregar respuestas finales.
Gold y Geng (2025)	Ciencia de datos	Tutor socrático basado en IA	Estudio sobre tutoría socrática	Analiza la utilidad de un tutor socrático zero-shot para guiar el aprendizaje mediante preguntas y apoyo progresivo.
Maalaoui y Samet (2026)	Ingeniería	Chatbot RAG	Diseño e implementación	Presenta un chatbot de generación aumentada por recuperación para estudiantes de primer año de ingeniería.
Lin et al. (2025)	STEM	Ayuda de aprendizaje basada en GPT	Estudio con analíticas multimodales	Analiza una herramienta GPT integrada con analíticas de aprendizaje para favorecer el aprendizaje autodirigido en STEM.

Categorías temáticas identificadas

A partir del análisis de los estudios seleccionados, se identificaron cuatro categorías principales: chatbots como tutores virtuales en STEM; retroalimentación personalizada y apoyo académico; aprendizaje autodirigido y pensamiento crítico; y riesgos, precisión y dependencia de la IA. Estas categorías permiten organizar la evidencia de manera sistemática y comprender que el uso de chatbots en educación superior STEM no debe analizarse solo como innovación tecnológica, sino como una estrategia pedagógica que requiere orientación, supervisión y diseño didáctico.

Tabla 2. Categorías temáticas derivadas del análisis

Categoría	Descripción	Número de estudios	Autores representativos
Chatbots como tutores virtuales en STEM	Agrupa estudios que analizan el uso de ChatGPT, LLM o chatbots como tutores académicos en asignaturas STEM.	4	Ding et al. (2023), Bassner et al. (2024), Caccavale et al. (2024), Maalaoui y Samet (2026)
Retroalimentación personalizada y apoyo académico	Incluye estudios que examinan la capacidad de los chatbots para ofrecer ayuda, guía, respuestas adaptadas y acompañamiento al estudiante.	3	Bassner et al. (2024), Gupta et al. (2025), Lin et al. (2025)
Aprendizaje autodirigido y pensamiento crítico	Reúne trabajos que relacionan el uso de IA con autonomía, autorregulación, razonamiento y aprendizaje activo.	3	Gold y Geng (2025), Lin et al. (2025), Ding et al. (2023)
Riesgos, precisión y dependencia de la IA	Considera estudios que advierten sobre errores conceptuales, respuestas incompletas, confianza excesiva y necesidad de mediación docente.	3	Ding et al. (2023), Gupta et al. (2025), Caccavale et al. (2024)

En relación con los chatbots como tutores virtuales en STEM, los estudios revisados muestran que estas herramientas pueden cumplir funciones de apoyo académico en asignaturas complejas. Ding et al. (2023) evidenciaron que ChatGPT puede ser percibido por los estudiantes como un tutor virtual útil en una clase universitaria de física, especialmente para resolver dudas y recibir explicaciones inmediatas. Bassner et al. (2024), por su parte, presentaron Iris como un tutor virtual orientado a ciencias de la computación, con capacidad para acompañar el desarrollo de ejercicios de programación. De manera similar, Caccavale et al. (2024) analizaron el papel de los modelos de lenguaje como tutores virtuales en ingeniería química, mientras que Maalaoui y Samet (2026) propusieron un chatbot RAG para apoyar a estudiantes de primer año de ingeniería. Estos hallazgos muestran que los chatbots pueden ampliar las oportunidades de acompañamiento académico, especialmente en contextos donde los estudiantes requieren apoyo frecuente para resolver problemas, revisar conceptos o avanzar a su propio ritmo.

Respecto a la retroalimentación personalizada y el apoyo académico, los estudios evidencian que los chatbots de IA pueden proporcionar respuestas inmediatas, explicaciones paso a paso y orientación adaptada a las necesidades del estudiante. Esta función resulta especialmente relevante en STEM, donde el aprendizaje suele requerir práctica constante, corrección de errores y comprensión progresiva de conceptos abstractos. Gupta et al. (2025) destacan que la evaluación de los modelos de lenguaje en tutoría matemática no debe limitarse a verificar si proporcionan la respuesta final correcta, sino que también debe considerar la calidad del razonamiento, la guía ofrecida y la capacidad para apoyar el proceso de resolución. En esta misma línea, Lin et al. (2025) muestran que las ayudas de aprendizaje basadas en GPT pueden integrarse con analíticas multimodales para favorecer procesos de aprendizaje autodirigido.

En cuanto al aprendizaje autodirigido y el pensamiento crítico, la evidencia sugiere que los chatbots pueden fortalecer la autonomía estudiantil cuando son utilizados como herramientas de acompañamiento y no como sustitutos del razonamiento propio. Gold y Geng (2025) aportan una perspectiva relevante al analizar un tutor socrático zero-shot, cuyo valor se centra en guiar al estudiante mediante preguntas, pistas y orientación progresiva. Este enfoque resulta importante porque evita que el chatbot funcione únicamente como generador de respuestas y lo aproxima más a una tutoría académica centrada en el razonamiento. De igual manera, Lin et al. (2025) relacionan el uso de herramientas GPT con el aprendizaje autodirigido en STEM, lo cual sugiere que la IA puede apoyar procesos de planificación, monitoreo y revisión del aprendizaje cuando se integra con criterios pedagógicos.

Sin embargo, los estudios también reportan riesgos y limitaciones. Uno de los principales desafíos es la precisión de las respuestas generadas por los modelos de lenguaje. En áreas STEM, donde los errores conceptuales, matemáticos o procedimentales pueden afectar la comprensión del estudiante, la confiabilidad de la información es un aspecto crítico. Ding et al. (2023) advierten que los estudiantes pueden confiar excesivamente en las respuestas de ChatGPT, incluso cuando estas no son completamente correctas. Gupta et al. (2025) también muestran la necesidad de evaluar la calidad de la tutoría matemática más allá del resultado final, debido a que una respuesta aparentemente correcta puede estar acompañada de razonamientos incompletos o poco pedagógicos. Por tanto, la evidencia sugiere que los chatbots deben ser utilizados con supervisión docente y con estrategias que promuevan la verificación, la reflexión y el pensamiento crítico.

De manera general, los resultados permiten afirmar que los chatbots basados en IA tienen potencial para actuar como tutores académicos en educación superior STEM, especialmente cuando ofrecen retroalimentación inmediata, apoyo personalizado, orientación paso a paso y recursos para el aprendizaje autónomo. No obstante, su efectividad depende del diseño pedagógico, del área disciplinar, de la calidad de las respuestas, de la capacidad del estudiante para evaluar críticamente la información y de la mediación docente. Por ello, los chatbots no deben entenderse como sustitutos del profesor, sino como herramientas complementarias para ampliar las oportunidades de aprendizaje.

Vacíos y tendencias emergentes

La revisión permitió identificar varios vacíos en la literatura. En primer lugar, aunque los estudios sobre chatbots y modelos de lenguaje en STEM han aumentado, todavía existe limitada evidencia empírica sobre su impacto real en el desempeño académico a largo plazo. Muchos trabajos se enfocan en percepciones, diseño de herramientas o evaluaciones iniciales, pero no siempre analizan si el uso del chatbot produce mejoras sostenidas en el aprendizaje.

En segundo lugar, se observa una necesidad de estudios comparativos entre distintas áreas STEM. La evidencia disponible muestra aplicaciones en física, programación, matemáticas, ingeniería química, ciencia de datos e ingeniería, pero todavía falta comprender si los chatbots funcionan de la misma manera en disciplinas con diferentes tipos de razonamiento, niveles de abstracción y formas de evaluación.

En tercer lugar, se identifica como tendencia emergente el diseño de chatbots más especializados, como tutores socráticos, sistemas RAG y herramientas GPT integradas con analíticas de aprendizaje. Esta evolución muestra un desplazamiento desde el uso general de ChatGPT hacia modelos más estructurados, contextualizados y orientados a objetivos pedagógicos específicos. Sin embargo, esta tendencia también plantea nuevos desafíos relacionados con la calidad de los datos, la transparencia de las respuestas, la privacidad, la integridad académica y la formación docente para el uso responsable de la IA.

DISCUSIÓN

Los resultados de esta revisión sistemática permiten responder al objetivo planteado al evidenciar que los chatbots basados en inteligencia artificial pueden actuar como tutores académicos en la educación superior STEM, siempre que su uso esté mediado por criterios pedagógicos claros, supervisión docente y actividades que promuevan el razonamiento del estudiante. Este hallazgo es relevante porque muestra que el valor de los chatbots no se encuentra únicamente en su capacidad para responder preguntas, sino en su potencial para acompañar procesos de aprendizaje complejos mediante retroalimentación inmediata, explicación de conceptos, guía paso a paso y apoyo personalizado.

En primer lugar, la evidencia revisada sugiere que los chatbots pueden ampliar las oportunidades de tutoría en asignaturas STEM. En áreas como física, programación, matemáticas, ingeniería química, ciencia de datos e ingeniería, los estudiantes suelen requerir apoyo frecuente para resolver problemas, comprender procedimientos y verificar razonamientos. En este contexto, los chatbots pueden funcionar como recursos de consulta permanente, disponibles fuera del horario de clase y capaces de responder de forma inmediata. Esta característica resulta especialmente útil en educación superior, donde la autonomía estudiantil es necesaria, pero no siempre suficiente para enfrentar contenidos de alta dificultad.

En segundo lugar, los hallazgos muestran que la retroalimentación personalizada es una de las funciones más prometedoras de los chatbots de IA. A diferencia de otros recursos digitales estáticos, los modelos conversacionales permiten establecer intercambios más flexibles con el estudiante, adaptando las respuestas a sus preguntas, errores o necesidades específicas. Sin embargo, la revisión también evidencia que la retroalimentación generada por IA debe ser evaluada con cautela. En STEM, una explicación incorrecta o un procedimiento incompleto puede generar aprendizajes equivocados. Por ello, el uso de chatbots como tutores académicos requiere estrategias de verificación, actividades de contraste y orientación docente para que el estudiante no acepte las respuestas de forma automática.

Otro resultado importante es que los chatbots pueden favorecer el aprendizaje autodirigido y el pensamiento crítico cuando su uso se orienta hacia la guía y no hacia la sustitución del esfuerzo cognitivo. Los estudios sobre tutores socráticos y herramientas GPT integradas con analíticas de aprendizaje muestran que la IA puede ayudar al estudiante a reflexionar, revisar pasos, formular preguntas y monitorear su proceso de aprendizaje. Esta evidencia sugiere que el chatbot es más valioso cuando promueve la construcción de conocimiento y no cuando se limita a entregar soluciones finales. En este sentido, la tutoría basada en IA debe diseñarse para estimular el razonamiento, la autonomía y la autorregulación, evitando que el estudiante dependa de la herramienta para resolver tareas sin comprenderlas.

En comparación con la literatura previa sobre chatbots educativos e inteligencia artificial generativa en educación superior, los hallazgos de esta revisión coinciden en reconocer beneficios como la personalización, la disponibilidad, la rapidez de respuesta y el apoyo al aprendizaje autónomo. Sin embargo, esta revisión también permite matizar el entusiasmo tecnológico, ya que en las áreas STEM la precisión, la argumentación y la trazabilidad del razonamiento son condiciones fundamentales. Por tanto, el uso de chatbots como tutores académicos no debe evaluarse únicamente por la satisfacción estudiantil o la facilidad de uso, sino también por su impacto en la comprensión conceptual, la resolución de problemas, la transferencia del aprendizaje y la calidad del pensamiento disciplinar.

Desde una perspectiva práctica, los resultados tienen implicaciones importantes para docentes e instituciones de educación superior. En primer lugar, se recomienda integrar los chatbots de IA como recursos complementarios dentro de actividades planificadas, evitando

su uso improvisado o exclusivamente instrumental. En segundo lugar, los docentes deben enseñar a los estudiantes a verificar respuestas, comparar fuentes, identificar errores y formular preguntas de calidad. En tercer lugar, las instituciones deben establecer lineamientos sobre integridad académica, privacidad, citación del uso de IA y límites del apoyo automatizado. Estas acciones son necesarias para que los chatbots contribuyan al aprendizaje sin debilitar el pensamiento crítico ni la responsabilidad académica.

En términos teóricos, esta revisión aporta al comprender los chatbots de IA como mediadores del aprendizaje en STEM y no solo como herramientas tecnológicas. Su función como tutores académicos depende de la interacción entre tres elementos: la capacidad técnica del modelo, el diseño pedagógico de la actividad y la competencia del estudiante para utilizar críticamente la herramienta. Desde el punto de vista metodológico, la revisión muestra la necesidad de investigaciones más robustas, especialmente estudios experimentales, longitudinales y comparativos que permitan evaluar el efecto real de los chatbots en el desempeño académico, la autorregulación, la comprensión conceptual y la resolución de problemas.

La presente revisión debe interpretarse considerando algunas limitaciones. En primer lugar, la muestra documental estuvo conformada por 7 estudios, lo que permite una síntesis focalizada, pero limita la generalización de los hallazgos. En segundo lugar, la rápida evolución de los modelos de lenguaje puede hacer que algunas evidencias pierdan vigencia en poco tiempo, debido a la aparición de nuevas versiones, mejoras técnicas y herramientas especializadas. En tercer lugar, los estudios seleccionados presentan diversidad de enfoques, áreas STEM y tipos de chatbot, lo que dificulta establecer comparaciones directas entre todos los resultados. Estas limitaciones no invalidan los hallazgos, pero sí sugieren la necesidad de continuar actualizando la evidencia sobre este campo emergente.

En conjunto, los hallazgos permiten sostener que los chatbots basados en inteligencia artificial tienen un potencial significativo como tutores académicos en educación superior STEM, pero su efectividad depende de una integración pedagógica responsable. La revisión contribuye al campo educativo al mostrar que estas herramientas pueden apoyar la retroalimentación, la tutoría, el aprendizaje autodirigido y la resolución de problemas, siempre que se acompañen de mediación docente, evaluación crítica y lineamientos institucionales. Futuras investigaciones deberían profundizar en el impacto real de los chatbots en el aprendizaje STEM, diferenciando entre usos generales de IA generativa y modelos de tutoría académica diseñados específicamente para objetivos educativos.

CONCLUSIONES

La presente revisión sistemática permitió analizar la producción científica reciente sobre el uso de chatbots basados en inteligencia artificial como tutores académicos en asignaturas STEM de educación superior, identificando sus aplicaciones pedagógicas, beneficios, limitaciones, riesgos y vacíos de investigación. En respuesta al objetivo planteado, los hallazgos evidenciaron que estas herramientas poseen un potencial significativo para apoyar el aprendizaje universitario mediante la retroalimentación inmediata, la explicación de conceptos, la orientación paso a paso, la personalización del apoyo académico y el fortalecimiento del aprendizaje autodirigido.

La evidencia analizada mostró que los chatbots pueden desempeñar funciones de tutoría en áreas como física, programación, matemáticas, ingeniería química, ciencia de datos e ingeniería. Sin embargo, su efectividad no depende únicamente de la capacidad tecnológica del modelo, sino de la forma en que se integra dentro del proceso educativo. Los mejores resultados se relacionan con usos pedagógicamente estructurados, como la tutoría socrática, los sistemas de recuperación aumentada por generación y las herramientas integradas con analíticas de aprendizaje.

A pesar de estas oportunidades, se identificaron riesgos asociados con respuestas incorrectas, razonamientos incompletos, dependencia excesiva, debilitamiento del esfuerzo cognitivo y problemas de integridad académica. Por ello, los chatbots no deben sustituir la mediación docente, sino complementarla mediante actividades que promuevan la verificación, el pensamiento crítico y la autonomía.

Se concluye que los chatbots de inteligencia artificial pueden constituirse en tutores académicos efectivos en educación superior STEM cuando su implementación responde a objetivos pedagógicos claros, supervisión docente y criterios éticos. Futuras investigaciones deberán evaluar su impacto longitudinal en el desempeño, la comprensión conceptual y la autorregulación del aprendizaje.

Referencias

Bassner, P., Frankford, E., & Krusche, S. (2024). Iris: An AI-driven virtual tutor for computer science education. *Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE*, 1, 394–400. https://doi.org/10.1145/3649217.3653543

Caccavale, F., Gargalo, C. L., Gernaey, K. V., & Krühne, U. (2024). Towards Education 4.0: The role of Large Language Models as virtual tutors in chemical engineering. *Education for Chemical Engineers, 49*, 1–11. https://doi.org/10.1016/j.ece.2024.07.002

Ding, L., Li, T., Jiang, S., & Gapud, A. (2023). *Students' perceptions of using ChatGPT in a physics class as a virtual tutor*. International Journal of Educational Technology in Higher Education, 20, Article 63. doi: 10.1186/s41239-023-00434-1

Gold, K., & Geng, S. (2025). On the helpfulness of a zero-shot Socratic tutor. *Lecture Notes on Data Engineering and Communications Technologies, 228*, 154–170. https://doi.org/10.1007/978-981-97-9255-9_11

Gupta, A., Reddig, J., Calò, T., Weitekamp, D., & MacLellan, C. (2025). Beyond final answers: Evaluating Large Language Models for math tutoring. *Lecture Notes in Computer Science, 15877 LNAI*, 323–337. https://doi.org/10.1007/978-3-031-98414-3_23

Lin, C.-J., Wang, W.-S., Lee, H.-Y., Li, P.-H., Huang, Y.-M., & Wu, T.-T. (2025). Advancing self-directed learning in STEM education: Integrating GPT-based learning aid with multimodal learning analytics. *Journal of Research on Technology in Education*. https://doi.org/10.1080/15391523.2025.2504358

Labadze, L., Grigolia, M., & Machaidze, L. (2023). *Role of AI chatbots in education: Systematic literature review*. International Journal of Educational Technology in Higher Education, 20, Article 56. doi: 10.1186/s41239-023-00426-1

Lee, H.-Y., Chen, P.-H., Wang, W.-S., Huang, Y.-M., & Wu, T.-T. (2024). *Empowering ChatGPT with guidance mechanism in blended learning: Effect of self-regulated learning, higher-order thinking skills, and knowledge construction*. International Journal of Educational Technology in Higher Education, 21, Article 16. doi: 10.1186/s41239-024-00447-4

Luan, L., Lin, X., & Dai, Y. (2025). *Bridging the gap: ChatGPT's role in enhancing STEM education*. Open Praxis, 17(1), 108–128. doi: 10.55982/openpraxis.17.1.685

Maalaoui, H., & Samet, Y. (2026). Use of a retrieval-augmented generation (RAG) chatbot for first-year engineering education at ESPRIT. *Lecture Notes of the Institute for Computer*

Sciences, Social-Informatics and Telecommunications Engineering, 676 LNICST, 438–449. https://doi.org/10.1007/978-3-032-16635-7_28

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71

Pepin, B., Buchholtz, N., & Salinas-Hernández, U. (2025). *A scoping survey of ChatGPT in mathematics education*. Digital Experiences in Mathematics Education, 11, 9–41. doi: 10.1007/s40751-025-00172-1

Sun, D., Boudouaia, A., Zhu, C., & Li, Y. (2024). *Would ChatGPT-facilitated programming mode impact college students' programming behaviors, performances, and perceptions? An empirical study*. International Journal of Educational Technology in Higher Education, 21, Article 14. doi: 10.1186/s41239-024-00446-5

Yusuf, A., Pervin, N., & Román-González, M. (2024). *Generative AI and the future of higher education: A threat to academic integrity or reformation? Evidence from multicultural perspectives*. International Journal of Educational Technology in Higher Education, 21, Article 21. doi: 10.1186/s41239-024-00453-6

ISSN: 3103-1315

SCIENTIA Research Club
ADVANCING RESEARCH, SHAPING MINDS.

LETTER OF ACCEPTANCE

CARTA DE ACEPTACIÓN DE ARTÍCULO CIENTÍFICO

Por la presente se certifica que el artículo titulado **"Chatbots basados en inteligencia artificial como tutores académicos en la educación superior STEM: una revisión sistemática de la literatura".**

Un trabajo de investigación de los autores: **Lenin Fernando Ortiz Pineda y Priscila Elizabeth Mackliff Montero** siendo su artículo revisado por Doble Par Ciego y Sistema de Doble Revisión Editorial, antes de ser publicado.

El artículo será publicado en la Revista Científica Research **Club Scientia** ISSN: 3103-1315, en la edición julio – septiembre, 2026, Volumen 4, Número 3. Verificable en nuestra plataforma: https://researchclubscientia.com/in dex.php/files

RESEARCH CLUB SCIENTIA

Ing. Luci Salas Narvaez, MSc.
EDITORA REVISTA RESEARCH CLUB SCIENTIA

T: +593 XX XXX XXXX
W: www.researchclubscientia.com
M: researchclubscientia@gmail.com

UNEMI

UNIVERSIDAD ESTATAL DE MILAGRO

¡Evolución académica!

@UNEMIEcuador