



Revista de Ciencias Sociales

Depósito legal ppi 201502ZU4662
Esta publicación científica en formato
digital es continuidad de la revista impresa
Depósito Legal: pp 197402ZU789
● ISSN: 1315-9518 ● ISSN-E: 2477-9431

Universidad del Zulia. Revista de la Facultad de Ciencias Económicas y Sociales
Vol. XXXII, No. 2

Abril – Junio 2026

Revista de Ciencias Sociales

Esta publicación científica en formato
digital es continuidad de la revista impresa
Depósito Legal: pp 197402ZU789
ISSN: 1315-9518

Inteligencia artificial generativa para la enseñanza de matemáticas y programación en educación superior: Revisión sistemática

Erazo-Arteaga, Víctor Alfonso*
Vacas Palacios, Santiago Marcelo**
Cisneros Ruales, Marcelo Bayardo***

Resumen

La rápida expansión de la inteligencia artificial generativa ha despertado un creciente interés en su aplicación educativa, especialmente en áreas que exigen razonamiento lógico y abstracción. El objetivo de esta revisión sistemática fue analizar el impacto de la inteligencia artificial generativa en el aprendizaje de estudiantes universitarios, así como en la preparación de clases por parte de los docentes. Se aplicaron las directrices PRISMA, definiendo criterios de inclusión y exclusión con base a preguntas PICO y PCC. La búsqueda en Scopus, Web of Science e IEEE Xplore permitió analizar 710 registros, de los cuales 67 cumplieron con los criterios establecidos. Los resultados muestran una literatura dominada por estudios descriptivos y con predominio de ChatGPT, donde se reportan beneficios en motivación y productividad, aunque sin evidencias concluyentes de mejoras en el rendimiento académico. La investigación se concentra en pocos países con ausencia de producción latinoamericana. En conclusión, el uso de inteligencia artificial generativa ofrece un potencial pedagógico prometedor, pero su efectividad depende de la integración crítica con la labor docente y de la generación de estudios experimentales más robustos que permitan establecer evidencias sólidas y generalizables.

Palabras clave: Inteligencia artificial generativa; ChatGPT; educación superior; enseñanza de matemáticas; enseñanza de programación.

* Magister Universitario en Métodos de Investigación en Educación. Magister en Diseño, Producción y Automatización. Ingeniero Mecánico. Docente e Investigador del Programa de Ingeniería Industrial en la Universidad Técnica del Norte, Ibarra, Imbabura, Ecuador. E-mail: vaerazo@utn.edu.ec ORCID: <https://orcid.org/0000-0001-5915-1864>

** Docente e Investigador del Programa de Ingeniería Industrial en la Universidad Técnica del Norte, Ibarra, Imbabura, Ecuador. E-mail: smvacas@utn.edu.ec ORCID: <https://orcid.org/0000-0003-3304-3843>

*** Docente e Investigador del Programa de Ingeniería Industrial en la Universidad Técnica del Norte, Ibarra, Imbabura, Ecuador. E-mail: mbcisneros@utn.edu.ec ORCID: <https://orcid.org/0000-0002-8818-6502>

Generative artificial intelligence for teaching mathematics and programming in higher education: A systematic review

Abstract

The rapid expansion of generative artificial intelligence has sparked growing interest in its educational applications, especially in areas requiring logical reasoning and abstraction. The objective of this systematic review was to analyze the impact of generative artificial intelligence on university student learning, as well as on lesson planning by faculty. The PRISMA guidelines were applied, defining inclusion and exclusion criteria based on PICO and PCC questions. A search of Scopus, Web of Science, and IEEE Xplore yielded 710 records, of which 67 met the established criteria. The results show literature dominated by descriptive studies, predominantly ChatGPT, which report benefits in motivation and productivity, although without conclusive evidence of improvements in academic performance. The research is concentrated in a few countries, with a lack of Latin American publications. In conclusion, the use of generative artificial intelligence offers promising pedagogical potential, but its effectiveness depends on its critical integration with teaching practices and the generation of more robust experimental studies that allow for the establishment of solid and generalizable evidence.

Keywords: Generative artificial intelligence; ChatGPT; higher education; mathematics education; programming education.

Introducción

La Investigación, el Desarrollo y la Innovación (I+D+I) han consolidado la tecnología como motor del progreso económico y social. Entre estos avances, la Inteligencia Artificial Generativa (GenAI), fruto de décadas de investigación, se destaca por su amplia aplicabilidad, condicionada únicamente por factores éticos y políticos (Al Naqbi et al., 2024; Cristofaro y Giardino, 2025; González et al., 2025).

La educación superior no es ajena a estos avances, diversos investigadores exploran la integración de la GenAI en la enseñanza, el aprendizaje, la docencia y la gestión académica (Mustafa et al., 2024; Perezchica-Vega et al., 2024). Sin embargo, esto genera debate sobre la efectividad en comparación con los métodos tradicionales de enseñanza (Sánchez-Ruiz y Larrea-Silva, 2022; Al Naqbi et al., 2024).

La matemática es una disciplina que fomenta el razonamiento y el análisis, competencias igualmente necesarias en la

programación, donde se requiere estructurar ideas en secuencias lógicas y tomar decisiones condicionales (Quiñones y Huiman, 2022; Carrasco-Barrionuevo et al., 2024; Morales et al., 2025). Las dificultades en matemáticas suelen trasladarse al aprendizaje de la programación, y viceversa; en este sentido, la GenAI puede apoyar en ambas áreas, al permitir personalizar la enseñanza, facilitar recursos de consulta permanente y monitorear la ética y el progreso estudiantil (Galindo, 2023; Karjanto, 2023; Lee, 2023; Yi et al., 2025).

La inserción de nuevas herramientas tecnológicas exige que su incorporación se sustente en evidencia científica. De esta manera, se evita que su incorporación responda a la novedad tecnológica o a tendencias pasajeras (Espina-Romero, 2022; Marín-González et al., 2025). Con el objetivo de analizar el impacto del uso de herramientas de inteligencia artificial generativa en el proceso de enseñanza-aprendizaje de las matemáticas y la programación en la educación superior,

se plantea este trabajo para generar una aproximación a su efecto en el desempeño de los estudiantes, la preparación de clases por parte de los docentes y la identificación de las herramientas GenAI más utilizadas en este ámbito. De este modo, se busca ofrecer un panorama actualizado que sirva de base para futuras decisiones pedagógicas y contribuciones académicas en el ámbito de la educación superior.

generativa en la preparación de clases y ejercicios por parte de docentes universitarios en las áreas de matemáticas y programación?; y, Q3: ¿Cuáles son las herramientas de inteligencia artificial generativa más utilizadas por docentes y estudiantes universitarios en la enseñanza y el aprendizaje de matemáticas y programación en el contexto de la educación superior?

1. Metodología

La revisión se desarrolló bajo las directrices PRISMA y se estructuró en tres fases: Construcción de la ecuación de búsqueda, codificación de datos y ejecución en bases indexadas. Las interrogantes planteadas fueron: Q1: ¿El uso de herramientas de inteligencia artificial generativa mejora el aprendizaje de las matemáticas y la programación en estudiantes universitarios comparado con métodos tradicionales de enseñanza?; Q2: ¿Cómo influye el uso de herramientas de inteligencia artificial

1.1. Construcción de la ecuación de búsqueda

Las preguntas de investigación fueron elaboradas utilizando las metodologías PICO (Paciente/Problema, Intervención, Comparación y *Outcome*-resultado) para la pregunta Q1; y, PCC (Población o participantes, Concepto y Contexto) para las preguntas Q2 y Q3. A partir de estas preguntas se identificaron componentes que ayudaron a deducir los criterios de inclusión y exclusión. El Cuadro 1, indica los componentes identificados, los criterios de inclusión y exclusión.

Cuadro 1
Criterios de inclusión y exclusión

Componentes	Criterio de inclusión	Criterio de exclusión
Docentes y estudiantes universitarios	Estudios que involucren docentes y/o estudiantes de educación superior.	Estudios centrados en educación primaria, secundaria o formación no universitaria.
Uso de inteligencia artificial generativa	Estudios que exploren, describan o evalúen herramientas de IA generativa.	Estudios que no se centren en IA o tecnologías no generativas.
Tipos de IA generativa más utilizadas	Investigaciones que identifiquen, herramientas de IA generativa utilizadas en el contexto educativo.	Estudios sin mención específica de herramientas IA o con tecnologías no aplicadas al ámbito educativo.
Métodos tradicionales de enseñanza	Estudios que comparen el uso de IA generativa con metodologías tradicionales de enseñanza.	Estudios donde no haya comparación. Otras revisiones sistemáticas.
Mejora del aprendizaje en matemáticas y programación	Estudios que evalúen el impacto de la IA generativa en el aprendizaje de matemáticas o programación en educación superior.	Estudios que se enfoquen en otras áreas disciplinarias.
Preparación de clases y ejercicios	Preparación de clases, materiales, ejercicios o recursos didácticos.	Estudios sin relación con la educación.
Enseñanza y aprendizaje de matemáticas y programación	Estudios que se enfoquen en la enseñanza y/o el aprendizaje de matemáticas, programación.	Estudios que no estén relacionados con la enseñanza

Fuente: Elaboración propia, 2026.

Con los términos de inclusión se establecieron los términos de búsqueda tanto en español como en inglés. Estos términos se organizaron mediante operadores booleanos,

combinando sinónimos con *OR* y componentes diferentes con *AND*, lo que permitió construir la ecuación de búsqueda principal. El Cuadro 2, indica los términos de búsqueda.

Cuadro 2
Términos de búsqueda

Componentes	Términos de búsqueda	Términos en inglés
Docentes y estudiantes universitarios	Educación superior	<i>Higher education</i>
	Estudiantes universitarios	<i>University students</i>
	Profesores universitarios	<i>College students</i>
	Instructores universitarios	<i>University teachers</i> <i>University instructors</i>
Uso de inteligencia artificial generativa	Inteligencia artificial generativa	<i>Generative artificial intelligence</i>
	IA generativa	<i>Generative AI</i>
	Modelos lingüísticos de gran tamaño	<i>Large language models</i>
	<i>ChatGPT</i>	<i>ChatGPT</i>
Enseñanza y aprendizaje de matemáticas y programación	Educación matemática	<i>Mathematics education</i>
	Educación en programación	<i>Programming education</i>
	Ciencias de la computación	<i>Computer science</i>
	Matemáticas	<i>Mathematics</i>
	Educación STEM	<i>STEM education</i>

Fuente: Elaboración propia, 2026.

A continuación, se detalla la ecuación primaria que se aplicó en las bases de datos: (*Higher education OR university students OR college students OR university teachers OR university instructors*) *AND* («*generative artificial intelligence*» *OR* «*generative AI*» *OR* «*large language models*» *OR* *ChatGPT*) *AND* («*mathematics education*» *OR* «*programming education*» *OR* «*computer science*» *OR* «*mathematics OR STEM education*»)

adaptaciones de instrumentos consolidados. Para los estudios descriptivos se empleó la *JBI Checklist* (Barker et al., 2022); mientras que los diseños cuasiexperimentales se valoraron con una versión ajustada de *ROBINS-I V2* (Cochrane Methods Group, 2024), y los ensayos experimentales con una adaptación de *RoB 2* (Sterne et al., 2019). Estas modificaciones tuvieron como propósito mantener criterios de validez interna, aplicabilidad y consistencia, favoreciendo la comparabilidad entre los diferentes tipos de estudios incluidos.

1.2. Codificación de datos

Se elaboró una matriz de codificación que organizó variables bibliográficas y contextuales de cada estudio, como año, país, institución, área de aplicación, carrera, herramienta GenAI utilizada, tipo de estudio y resultados principales. Esta sistematización permitió aplicar de forma coherente los criterios de inclusión y exclusión definidos en la revisión.

El riesgo de sesgo se evaluó mediante

1.3. Ejecución en bases indexadas

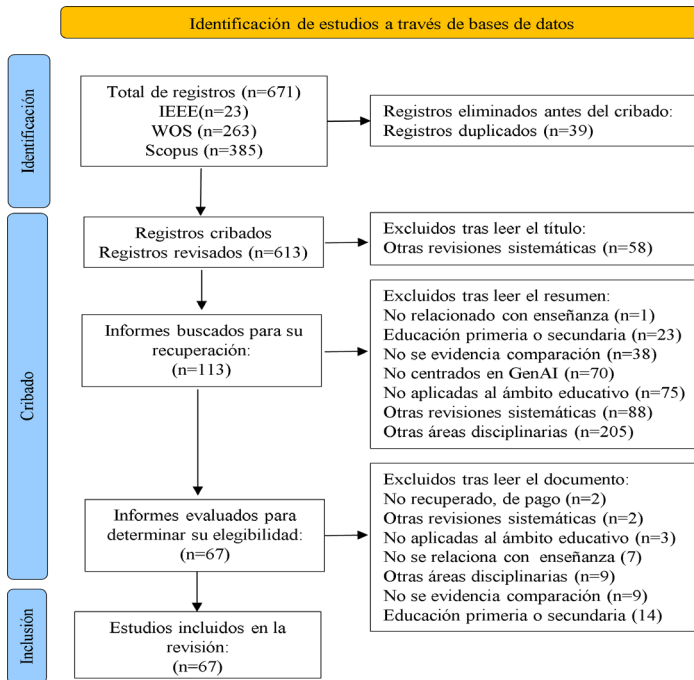
La búsqueda bibliográfica se realizó en marzo de 2025 en *Scopus*, *Web of Science* (WoS) e *IEEE Xplore*. Se incluyeron artículos publicados entre enero de 2014 y diciembre de 2024, en inglés y español, revisados por pares y de acceso abierto. Se excluyó literatura gris y documentos no arbitrados con el fin de preservar la validez y la transparencia de la

evidencia recopilada.

Los registros fueron exportados con metadatos de citación y contenido descriptivo, lo que facilitó la aplicación de los criterios de inclusión y exclusión. En total, se identificaron 710 registros; tras eliminar duplicados y aplicar el cribado, se descartaron 113 artículos en texto completo.

Con la matriz de codificación y la adaptación de las herramientas de evaluación de sesgo (Robleda, 2019; Wilson, 2019),

el primer y segundo autor revisaron de manera independiente los 113 artículos, contrastándolos con los criterios de inclusión y exclusión. Posteriormente, en reuniones de trabajo, los tres autores discutieron las discrepancias y, por consenso, se determinó la inclusión final de 67 estudios. La Figura 1, muestra el procedimiento de identificación, cribado, elegibilidad e inclusión de los estudios.



Fuente: Elaboración propia, 2026 a partir del método PRISMA.

Figura 1: Diagrama de flujo del proceso de selección de estudios

2. Resultados y discusión

Se incorporaron 3 artículos provenientes de la base de datos IEEE, 27 de *Scopus* y 37 de WoS. En relación con el periodo de publicación, se identificaron 2 artículos correspondientes al

año 2022, 17 al año 2023, y 47 al año 2024. En cuanto al diseño metodológico, 54 artículos son de carácter descriptivo, 11 presentan un diseño cuasiexperimental, y 2 corresponden a un diseño experimental. El conjunto de artículos seleccionados se sintetiza en el Cuadro 3.

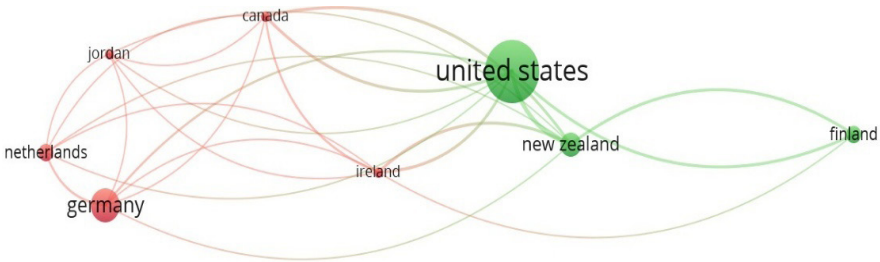
Cuadro 3 Artículos científicos incluidos

ID	Referencia	ID	Referencia
Doc. 1	Dingle y Krulis (2024)	Doc. 35	Sun et al. (2024)
Doc. 2	Frankford et al. (2024)	Doc. 36	Jošt et al. (2024)
Doc. 3	Xue et al. (2024)	Doc. 37	Humble et al. (2023)
Doc. 4	Phung et al. (2024)	Doc. 38	He et al. (2024)
Doc. 5	Abdulla et al. (2024)	Doc. 39	Segal y Biton (2024)
Doc. 6	Kazemitabaar et al. (2024)	Doc. 40	Rajala et al. (2023)
Doc. 7	Jordan et al. (2024)	Doc. 41	Singh et al. (2023)
Doc. 8	Kimmel et al. (2024)	Doc. 42	Mendonça (2024)
Doc. 9	Prather et al. (2023)	Doc. 43	Udias et al. (2024)
Doc. 10	Dasari et al. (2023)	Doc. 44	Getenet (2024)
Doc. 11	Dunder et al. (2024)	Doc. 45	Kosar et al. (2024)
Doc. 12	Husain (2024)	Doc. 46	Pabreja y Pabreja (2024)
Doc. 13	Bikanga (2024)	Doc. 47	López-Fernández y Vergaz (2024)
Doc. 14	X. Hou et al. (2024)	Doc. 48	Logacheva et al. (2024)
Doc. 15	Denny, Prather et al. (2024)	Doc. 49	Da Silva et al. (2024)
Doc. 16	Cipriano y Alves (2024)	Doc. 50	Gouia-Zarrad y Gunn (2024)
Doc. 17	Delima et al. (2024)	Doc. 51	Roest et al. (2023)
Doc. 18	Budhiraja et al. (2024)	Doc. 52	Montella et al. (2024)
Doc. 19	Joshi et al. (2024)	Doc. 53	Padiyath et al. (2024)
Doc. 20	MacNeil et al. (2024)	Doc. 54	Sánchez-Ruiz et al. (2023)
Doc. 21	I. Hou et al. (2024)	Doc. 55	Denny, Leinonen et al. (2024)
Doc. 22	Wahba et al. (2024)	Doc. 56	Meissner et al. (2024)
Doc. 23	Savelka, Agarwal, et al. (2023)	Doc. 57	Jacobs y Jaschke (2024)
Doc. 24	Sarsa et al. (2022)	Doc. 58	Shen et al. (2024)
Doc. 25	Zastudil et al. (2023)	Doc. 59	Doughty et al. (2024)
Doc. 26	Liffiton et al. (2023)	Doc. 60	Ouh et al. (2023)
Doc. 27	Berrezueta-Guzman y Krusche (2023)	Doc. 61	Neumann et al. (2025)
Doc. 28	Jin y Kim (2023)	Doc. 62	Karjanto (2023)
Doc. 29	Savelka, Ashley et al. (2023)	Doc. 63	Pankiewicz y Baker (2024)
Doc. 30	Haindl y Weinberger (2024)	Doc. 64	Cipriano y Alves (2023)
Doc. 31	Liao et al. (2024)	Doc. 65	Birillo et al. (2024)
Doc. 32	Rasnayaka et al. (2024)	Doc. 66	Azaiz et al. (2024)
Doc. 33	Dilling y Herrmann (2024)	Doc. 67	Azaiz et al. (2023)
Doc. 34	Richards et al. (2024)		

Fuente: Elaboración propia, 2026.

Los países con mayor número de aportes en la temática analizada muestran una marcada concentración. Estados Unidos encabeza con 18 contribuciones, seguido por Alemania (9), Nueva Zelanda (7), Países Bajos (5), y Finlandia (4). Con 3 publicaciones aparecen Portugal, Reino Unido, Canadá y Jordania.

La Figura II, evidencia la red de coautorías entre países, con Estados Unidos como nodo central y Alemania como un segundo eje de interacción. Otros países, como Canadá y Finlandia, muestran vínculos más limitados y colaboraciones esporádicas con naciones como Irlanda.

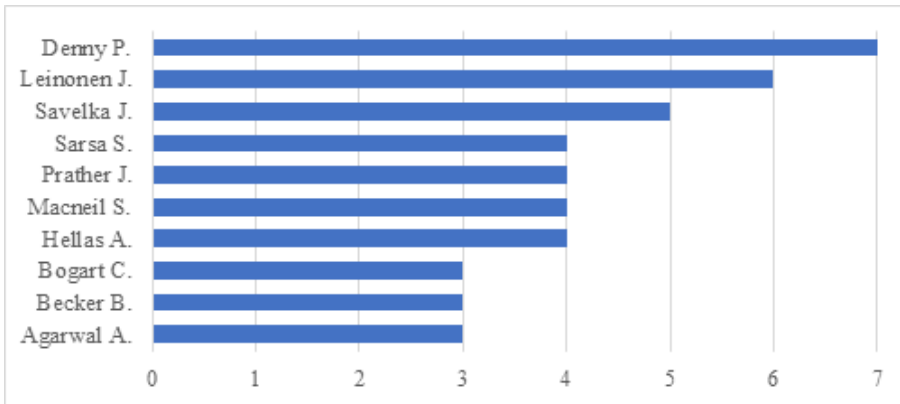


Fuente: Elaboración propia, 2026.

Figura II: Coautorías entre países en las publicaciones analizadas

El Gráfico I, muestra la distribución de documentos según autor. Se observa que Denny P. es el autor con mayor productividad,

seguido por Leinonen J. y Savelka J. Otros autores presentan 4 y 3 publicaciones.

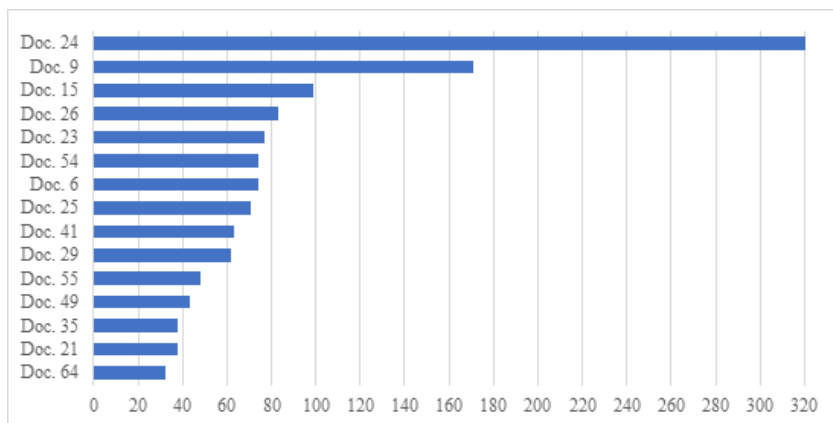


Fuente: Elaboración propia, 2026.

Gráfico I: Productividad de los autores según número de publicaciones

El Gráfico II, presenta los documentos más citados. El documento 24 destaca con más de 300 citas, seguido por el documento 9 con

170 citas, y el documento 15 con alrededor de 100. Un grupo intermedio de trabajos acumula entre 65 y 80 citas.



Fuente: Elaboración propia, 2026.

Gráfico II: Cantidad de citas recibidas por documento

El Cuadro 4, presenta las principales fuentes en las que se encontraron documentos incluidos en la revisión sistemática. La

fuente con mayor número de aportes es *ACM International Conference Proceeding Series*.

Cuadro 4

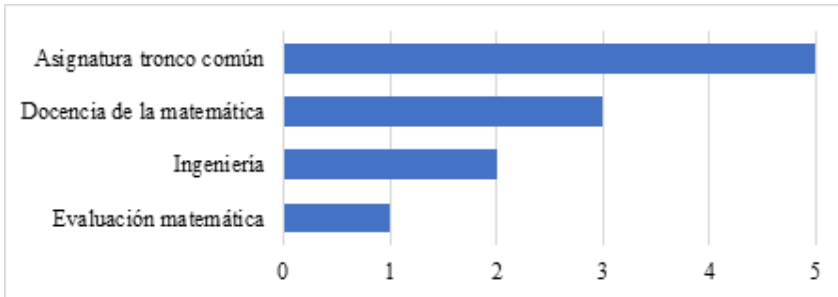
Fuentes de publicación con mayor número de documentos identificados

Fuentes	Documentos
ACM International Conference Proceeding Series	6
Applied Sciences-Basel	3
Education Sciences	3
International Electronic Journal of Mathematics Education	3
Proceedings - International Conference on Software Engineering	3
2024 IEEE Global Engineering Education Conference, EDUCON 2024	2
20th Annual ACM Conference on International Computing Education Research, ICER 2024, Vol 1	2
ACM Transactions on Computing Education	2
Conference on Human Factors in Computing Systems - Proceedings	2
Electronic Journal of E-Learning	2

Fuente: Elaboración propia, 2026.

Como se indica en el Gráfico III, se identificaron 11 estudios en matemáticas y 56 en programación. En matemáticas, cinco corresponden a asignaturas de tronco común

(álgebra, cálculo, probabilidad, estadística y geometría), tres a formación docente, dos a cursos de ingeniería, y, uno a evaluación matemática.

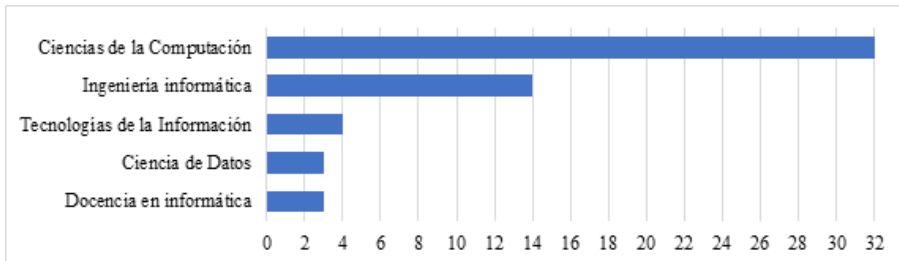


Fuente: Elaboración propia, 2026.

Gráfico III: Distribución de estudios en el área de la matemática

Según se indica en el Gráfico IV, 32 estudios se enfocan en Ciencias de la Computación, abarcando desde cursos introductorios hasta estructuras de datos y depuración de errores. Catorce (14) corresponden a Ingeniería Informática, con énfasis en inteligencia artificial, *software* y

modelos de lenguaje. Además, se registraron cuatro (4) estudios en Tecnologías de la Información, tres (3) en Ciencia de Datos, y tres (3) en formación docente en informática. Esta distribución refleja un mayor interés en la enseñanza de programación, frente a la menor atención recibida por las matemáticas.



Fuente: Elaboración propia.

Gráfico IV: Distribución de estudios en el área de programación

2.1. RQ1: La GenAI en el aprendizaje de las matemáticas

En el área de matemáticas se identificaron dos estudios cuasiexperimentales y nueve descriptivos. Los documentos 10 y 22 corresponden a los cuasiexperimentales, cuyo análisis no muestra indicios claros de sesgo, aunque tampoco permite asegurar que estén libres de él. El documento 10 evaluó tres condiciones: Uso exclusivo de GenAI, uso de GenAI con apoyo docente y

enseñanza tradicional. El grupo que usó solo GenAI obtuvo el peor desempeño; mientras que la combinación con el docente alcanzó los mejores resultados, sin diferencias significativas respecto a la enseñanza tradicional. El documento 22 aplicó un diseño *pretest-postest* en un curso de estadística, comparando un grupo con enseñanza tradicional frente a otro con GenAI. Los resultados muestran un impacto positivo en el razonamiento estadístico y en las actitudes hacia la asignatura.

Estos hallazgos no son suficientes para sostener que la GenAI mejore el aprendizaje matemático. Ambos estudios tienen muestras pequeñas (29 estudiantes en el Doc. 10 y 56 en el Doc. 22) y limitaciones en el diseño. Además, los dos estudios advierten sobre riesgos de usar GenAI como sustituto del docente.

En las investigaciones descriptivas, los documentos 33, 44, 50 y 62, presentan bajo riesgo de sesgo; mientras que los documentos 17, 39, 43, 54 y 56, muestran un riesgo moderado. Los primeros, ofrecen resultados con mayor consistencia metodológica; mientras que los segundos, presentan limitaciones que reducen la confianza en los hallazgos, aunque no los invalidan.

El documento 33 evidenció en los estudiantes ideas erróneas sobre el funcionamiento de la GenAI, escasa experiencia en su uso y una tendencia a emplear estrategias simples, como copiar y pegar los ejercicios a resolver. El documento 44 comparó las estrategias de resolución de los estudiantes con las de la GenAI; se observó similitudes, pero también errores frecuentes y poca fiabilidad. El documento 50 mostró que el uso de la GenAI aumentó el interés, el compromiso y las habilidades, además los estudiantes recomendaron incorporar la GenAI en la enseñanza. El documento 62 describe errores en álgebra lineal con apoyo de *SageMath* y GenAI, principalmente en diagonalización y descomposición, que fueron interpretados como oportunidades para estimular el pensamiento crítico.

En los estudios de riesgo moderado se usaron encuestas, análisis comparativos y valoraciones de expertos. El documento 17, con escalas tipo *Likert*, mostró un efecto positivo de la GenAI en el aprendizaje autorregulado, pero no en la ansiedad matemática ni en el rendimiento académico. El documento 39 afirma que la GenAI estimula el conocimiento, el pensamiento crítico y creativo. Finalmente, el documento 54 reportó que la GenAI fue aceptada como herramienta de apoyo por estudiantes de ingeniería, aunque se señalaron problemas de fiabilidad

en cálculos y posibles efectos negativos en competencias transversales.

En conjunto, los estudios descriptivos muestran que la GenAI puede favorecer la motivación, el compromiso y la autorregulación del aprendizaje. No obstante, también se reportan efectos contrarios vinculados a la inexperiencia de los estudiantes y a los errores de la herramienta, lo que plantea dudas sobre su impacto en competencias como el pensamiento analítico y el trabajo en equipo.

2.2. RQ1: La GenAI en el aprendizaje de la programación

En la enseñanza de la programación se identificaron dos estudios experimentales, nueve cuasiexperimentales, y cuarenta y cinco descriptivos. Los documentos 3 y 45 corresponden a los experimentales, que presentan pequeñas fuentes de sesgo, aunque no alcanzan a considerarse sesgados, por lo que sus hallazgos deben interpretarse con cautela.

El documento 3 se llevó a cabo en un curso introductorio de programación en grupos de control y experimental. No se halló diferencia significativa entre los estudiantes que usaron GenAI y los que no. El uso de la GenAI no garantizó un mejor aprendizaje ni un peor desempeño. Se observó que varios estudiantes generaron dependencia a la GenAI y la mayoría tuvo actitud neutral o ligeramente positiva hacia ella. Se expresó preocupación sobre fiabilidad y ética. El documento 45 evaluó el impacto de la GenAI en el aprendizaje mediante una comparación directa entre dos grupos. No hubo diferencias significativas entre los grupos en cuanto a rendimiento, esto podría indicar que, bajo planificación, el uso de la GenAI no afecta negativamente ni mejora significativamente el aprendizaje.

Los estudios cuasiexperimentales presentan limitaciones metodológicas, aunque aportan tendencias relevantes. Los documentos 5, 31 y 35, indican mejoras en el rendimiento y en el pensamiento computacional, aunque

los efectos positivos no siempre fueron consistentes ni significativos. Los documentos 14, 40 y 63, evidencian beneficios percibidos en comprensión, productividad y actitudes hacia la GenAI, pero también se reportan riesgos asociados, como la disminución de habilidades de resolución de problemas (Doc. 31) o una correlación negativa entre el uso intensivo y los resultados académicos (Doc. 36). En algunos casos, el desempeño de la AI superó al de los estudiantes en tareas específicas, como la detección de errores de código (Doc. 20), lo que plantea interrogantes sobre su papel como herramienta de aprendizaje.

En los estudios descriptivos con calidad alta, el documento 61 muestra una alta aceptación de la GenAI por su facilidad de uso y utilidad percibida. El documento 53, por su parte, señala que el uso de la GenAI está mediado por percepciones sociales y que la dependencia percibida puede reducir la autoeficacia y el rendimiento. Ambos estudios advierten riesgos vinculados a la dependencia.

En los estudios descriptivos con calidad moderada se observa varios enfoques, desde evaluaciones comparativas hasta percepciones de estudiantes y docentes. Los hallazgos muestran que la GenAI puede resolver tareas y generar retroalimentación con niveles de precisión variables, destacando buenos resultados en problemas básicos (Docs. 1, 4, 15, 42, 59), pero limitaciones claras en actividades más complejas (Docs. 11, 29, 64). También se reportan valoraciones positivas de estudiantes y profesores en cuanto a utilidad, productividad y motivación (Docs. 6, 12, 30, 32, 46, 49, 55, 57, 65), aunque acompañadas de preocupaciones recurrentes sobre la calidad de los resultados y desarrollo de habilidades críticas. Los documentos 13, 41 y 47, señalan que el uso de la GenAI está condicionado por factores sociales y por el desconocimiento de estrategias adecuadas de uso.

En los estudios descriptivos con riesgo de sesgo alto, los resultados se dividen entre observaciones críticas y valoraciones positivas. Algunos trabajos advierten problemas de fiabilidad, superficialidad en las respuestas

o incumplimiento de buenas prácticas de programación (Docs. 16, 19, 23, 38). A la vez, se recogen percepciones positivas de estudiantes y docentes respecto a la utilidad y comodidad de estas herramientas (Docs. 18, 21, 25, 26, 27, 37), aunque acompañadas de preocupaciones sobre plagio y limitaciones en el desarrollo de habilidades críticas.

2.3. RQ2: La GenAI en la preparación de clases y ejercicios de matemáticas

Los trabajos 9, 28, 48 y 60, destacan el potencial de la GenAI para generar materiales, y apoyar procesos de tutoría; mientras que los documentos 50 y 54, describen su uso en la planificación de ejercicios de refuerzo y en la personalización de tareas. Aunque los documentos 51, 52, 66 y 67, ratifican estos avances, también advierten que los errores detectados limitan su fiabilidad.

El documento 62 refuerza esta advertencia al señalar que, si bien la GenAI puede automatizar cálculos rutinarios, la presencia de imprecisiones exige mantener una supervisión constante. Así mismo, pese a que el documento 22 señala que la herramienta ofrece retroalimentación personalizada y facilita la adaptación de la enseñanza a necesidades individuales. El documento 10 señala que, la efectividad en la comprensión de contenidos depende de la claridad de las preguntas formuladas.

Las limitaciones en la precisión de la herramienta también se evidencian en el documento 43, que al comparar exámenes de 608 estudiantes con respuestas generadas por GenAI, encontró que estas superaron el promedio en estadística, pero fueron inferiores en álgebra. Además, aunque ofrecieron explicaciones más detalladas, estas no siempre resultaron correctas. En la misma dirección, el documento 56 evaluó 240 ítems producidos con GenAI y mostró que, aunque la herramienta puede alinearse con competencias y niveles cognitivos, comete errores que requieren revisión experta.

2.4. RQ2: La GenAI en la preparación de clases y ejercicios de programación

En los documentos 7, 24, 28 y 59, se reporta el uso de la GenAI para generar ejercicios, materiales de clase, explicaciones, bancos de preguntas y pruebas automatizadas, con resultados comparables a los producidos por los docentes. A su vez, los documentos 23, 27, 29 y 58, plantean estrategias para el diseño de tareas que incorporan enunciados más contextualizados, vinculados a los recursos del curso y a los proyectos desarrollados.

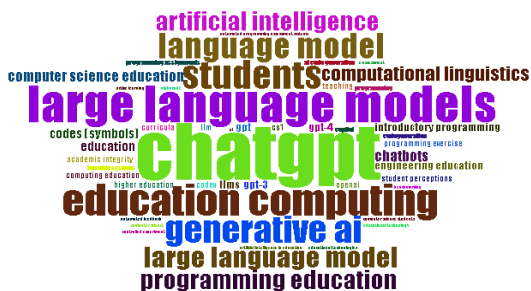
La herramienta también ha sido integrada en la retroalimentación del código, como señalan los documentos 2, 4, 31, 35, 51, 57, 65 y 67, mediante pistas personalizadas o comentarios específicos que permiten detectar errores, sugerir siguientes pasos y acompañar la resolución de problemas. De manera complementaria, los documentos 1, 15, 20, 31 y 66, describen su uso para mejorar mensajes de error, sintetizar recomendaciones pedagógicas y automatizar parcialmente la corrección.

La GenAI permitiría gestionar consultas fuera de clase, proveer apoyo continuo, y reducir carga operativa en la elaboración y revisión de materiales (Docs. 1, 11, 61). Algunas experiencias describen paneles y reportes semanales que permiten al docente monitorizar uso, identificar dificultades y ajustar la enseñanza durante el semestre (Doc. 6).

Los documentos 16, 23, 27, 45 y 58, proponen orientaciones para incorporar la GenAI en clase, promoviendo la evaluación crítica de las respuestas, la defensa de tareas, el trabajo en papel, y el diseño de ejercicios que limiten respuestas triviales de la herramienta. Los documentos 55 y 58, sugieren que se podría integrar ingeniería de *prompts* y actividades tipo *Prompt Problems* en el currículo. En cuanto a esto, los documentos 7, 38, 41, 48, 52, 60, 62 y 66, señalan la necesidad de la supervisión docente ante errores y explicaciones inconsistentes así también la implementación de políticas institucionales sobre uso responsable de la GenAI.

2.5. RQ3: Herramientas de inteligencia artificial generativa más utilizadas

La Figura III, muestra que los términos más recurrentes en los estudios revisados son “*ChatGPT*”, “*large language models*” y “*education computing*”, lo que confirma la centralidad de *ChatGPT* en la adopción de GenAI en educación. Tal como sugiere Lee (2023), aunque no fue la primera herramienta de inteligencia artificial generativa, sí fue la primera en alcanzar una adopción masiva y global en un periodo muy breve tras su lanzamiento en noviembre de 2022 (Adamopoulou y Moussiades, 2020; Lee, 2023).

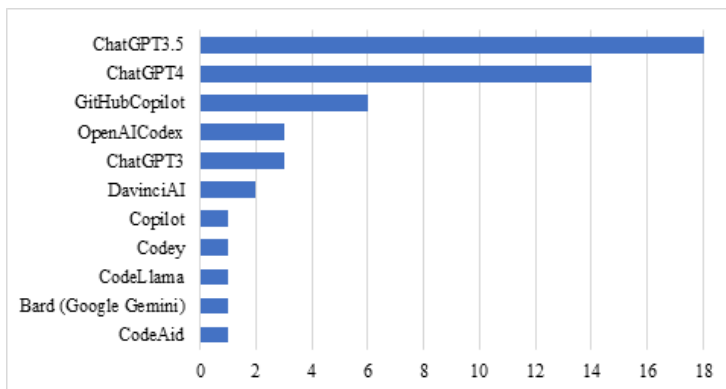


Fuente: Elaboración propia, 2026.

Figura III: Palabras clave en los estudios revisados

El Gráfico V, evidencia el uso de *ChatGPT*, especialmente en sus versiones 3.5 y 4, esto la consolida como la herramienta más usada. No obstante, también se reportan otras opciones como *Copilot*, *Codex* y *Genini*, lo

que muestra que, aunque el interés principal se centra en *ChatGPT*, la investigación también explora alternativas y variantes, en ocasiones combinadas en un mismo estudio.



Fuente: Elaboración propia, 2026.

Gráfico V: Herramientas de IA mencionadas en los estudios revisados

En esta revisión no se realizó un análisis formal del sesgo de publicación; la heterogeneidad metodológica de los estudios incluidos, el predominio de diseños descriptivos y la restricción a artículos de acceso abierto no permitieron aplicar pruebas estadísticas tradicionales como *funnel plot* o *test de Egger*. Aunque se emplearon tres bases de datos indexadas y se consideraron estudios en inglés y español, no puede descartarse la existencia de sesgo de publicación, en particular por la concentración geográfica de las investigaciones. Esta limitación debe tenerse en cuenta al interpretar los resultados y refuerza la necesidad de ampliar la búsqueda en futuras revisiones hacia bases de datos adicionales e incluir artículos de acceso restringido.

Conclusiones

El predominio de estudios descriptivos basados en datos subjetivos, la escasez de

investigaciones con diseños experimentales, el tamaño reducido de las muestras en estudios cuasiexperimentales y los riesgos de sesgo identificados limitan la reproducibilidad y la capacidad de generalizar los hallazgos sobre el uso de la GenAI en matemáticas y programación. Desde la perspectiva del diseño experimental, esta situación sugiere que los resultados actuales constituyen más bien indicios exploratorios que evidencias concluyentes.

En este contexto, los hallazgos deben interpretarse con cautela, puesto que la percepción de los estudiantes respecto a una mejora en su aprendizaje no necesariamente se traduce en un aprendizaje significativo alineado con los objetivos formativos. Es posible que los estudiantes se sientan más competentes tras usar GenAI, cuando en realidad ello no implica necesariamente la adquisición de un conocimiento profundo o duradero. De hecho, varios autores coinciden en que el uso de inteligencia artificial generativa en matemáticas y programación

no garantiza mejoras en el aprendizaje, y que su mayor efectividad se alcanza cuando se integra de manera complementaria a la guía del docente.

La investigación sobre GenAI en la enseñanza de matemáticas y programación se concentra en un número limitado de países y editoriales. La ausencia de participación latinoamericana en estas dinámicas no solo reduce la visibilidad internacional de la región, sino que también excluye contextos socioculturales y educativos con particularidades propias. Esta situación indica la necesidad de fortalecer redes de colaboración y consolidar espacios editoriales en América Latina, de modo que sus aportes puedan integrarse de manera equilibrada al debate académico sobre la incorporación de nuevas tecnologías en la educación superior.

Referencias bibliográficas

- Abdulla, S., Ismail, S., Fawzy, Y., y Elhaj, A. (2024). Using ChatGPT in Teaching Computer Programming and Studying its Impact on Students Performance. *Electronic Journal of E-Learning*, 22(6), 66-81. <https://doi.org/10.34190/EJEL.22.6.3380>
- Adamopoulou, E., y Moussiades, L. (2020). Chatbots: History, technology, and applications. *Machine Learning with Applications*, 2, 100006. <https://doi.org/10.1016/j.mlwa.2020.100006>
- Al Naqbi, H., Bahroun, Z., y Ahmed, V. (2024). Enhancing work productivity through generative artificial intelligence: A comprehensive literature review. *Sustainability*, 16(3), 1166. <https://doi.org/10.3390/su16031166>
- Azaiz, I., Deckarm, O., y Strickroth, S. (2023). AI-Enhanced auto-correction of programming exercises: How effective is GPT-3.5? *International Journal of Engineering Pedagogy*, 13(8), 67-83. <https://doi.org/10.3991/ijep.v13i8.45621>
- Azaiz, I., Kiesler, N., y Strickroth, S. (2024). Feedback-Generation for programming exercises with GPT-4. *Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE, 1*, 31-37. <https://doi.org/10.1145/3649217.3653594>
- Barker, T. H., Stone, J. C., Sears, K., Klugar, M., Leonardi-Bee, J., Tufanaru, C., Aromataris, E., y Munn, Z. (2022). Revising the JBI quantitative critical appraisal tools to improve their applicability: An overview of methods and the development process. *JBI Evidence Synthesis*, 21(3), 478-493. <https://doi.org/10.11124/JBIES-22-00125>
- Berrezueta-Guzman, J., y Krusche, S. (2023). Recommendations to Create Programming Exercises to Overcome ChatGPT. *IEEE 35th International Conference on Software Engineering Education and Training (CSEET&T)*, Tokyo, Japan, 2023, pp. 147-151. <https://doi.org/10.1109/CSEET58097.2023.00031>
- Bikanga, M. (2024). It helps with crap lecturers and their low effort: Investigating computer science students' perceptions of using ChatGPT for learning. *Education Sciences*, 14(10), 1106. <https://doi.org/10.3390/educsci14101106>
- Birillo, A., Artser, E., Potriasaeva, A., Vlasov, I., Dziales, K., Golubev, Y., Gerasimov, I., Keuning, H., y Bryksin, T. (2024). One step at a time: Combining LLMs and static analysis to generate next-step hints for programming tasks. *Proceedings of the 24th Koli Calling International Conference on Computing Education Research*, 9. <https://doi.org/10.1145/3699538.3699556>
- Budhiraja, R., Joshi, I., Challa, J. S.,

- Akolekar, H. D., y Kumar, D. (2024). "It's not like Jarvis, but it's pretty close!" - Examining ChatGPT's usage among undergraduate students in computer science. *ACM International Conference Proceeding Series*, pp. 124-133. <https://doi.org/10.1145/3636243.3636257>
- Carrasco-Barrionuevo, J. W., Allauca-Pancho, F. R., Mantilla-Cabrera, C. E., y Santillán-Lima, J. C. (2024). Estrategia didáctica para el aprendizaje de cálculo en estudiantes universitarios con el software Derive. *Revista de Ciencias Sociales (Ve)*, XXX(E-10), 222-242. <https://doi.org/10.31876/rcs.v30i.42840>
- Cipriano, B. P., y Alves, P. (2023). GPT-3 vs Object Oriented Programming Assignments: An experience report. *Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE, 1*, 61-67. <https://doi.org/10.1145/3587102.3588814>
- Cipriano, B. P., y Alves, P. (2024). LLMs Still Can't Avoid Instanceof: An Investigation Into GPT-3.5, GPT-4 and Bard's Capacity to Handle Object-Oriented Programming Assignments. *Proceedings - International Conference on Software Engineering*, 162-169. <https://doi.org/10.1145/3639474.3640052>
- Cochrane Methods Group (2024). *The Risk of Bias In Non-randomized Studies – of Interventions, Version 2 (ROBINS-I V2) assessment tool (for follow-up studies)*. ROBINS-I Development Group. <https://www.riskofbias.info/welcome/robins-i-v2>
- Cristofaro, M., y Giardino, P. L. (2025). Surfing the AI waves: the historical evolution of artificial intelligence in management and organizational studies and practices. *Journal of Management History*. <https://doi.org/10.1108/JMH-01-2025-0002>
- Dasari, D., Hendriyanto, A., Sahara, S., Suryadi, D., Muhaimin, L. H., Chao, T., y Fitriana, L. (2023). ChatGPT in didactical tetrahedron, does it make an exception? A case study in mathematics teaching and learning. *Frontiers in Education*, 8, 1295413. <https://doi.org/10.3389/educ.2023.1295413>
- Da Silva, C. A. G., Ramos, F. N., De Moraes, R. V., y Dos Santos, E. L. (2024). ChatGPT: Challenges and benefits in software programming for higher education. *Sustainability*, 16(3), 1245. <https://doi.org/10.3390/su16031245>
- Delima, N., Kusuma, D. A., y Paulus, E. (2024). The students' mathematics self-regulated learning and mathematics anxiety based on the use of ChatGPT, music, study program, and academic achievement. *Infinity Journal*, 13(2), 349-362. <https://doi.org/10.22460/infinity.v13i2.p349-362>
- Denny, P., Leinonen, J., Prather, J., Luxton-Reilly, A., Amarouche, T., Becker, B. A., y Reeves, B. N. (2024). Prompt Problems: A New Programming Exercise for the Generative AI Era. *SIGCSE 2024 - Proceedings of the 55th ACM Technical Symposium on Computer Science Education, 1*, 296-302. <https://doi.org/10.1145/3626252.3630909>
- Denny, P., Prather, J., Becker, B. A., Finnie-Ansley, J., Hellas, A., Leinonen, J., Luxton-Reilly, A., Reeves, B. N., Santos, E. A., y Sarsa, S. (2024). Computing Education in the Era of Generative AI. *Communications of the ACM*, 67(2), 56-67. <https://doi.org/10.1145/3624720>
- Dilling, F., y Herrmann, M. (2024). Using large language models to support pre-service teachers mathematical reasoning—an exploratory study on ChatGPT as an instrument for creating

- mathematical proofs in geometry. *Frontiers in Artificial Intelligence*, 7, 1460337. <https://doi.org/10.3389/frai.2024.1460337>
- Dingle, A., y Krulis, M. (2024). Tackling students' coding assignments with LLMs. *Proceedings - 2024 International Workshop on Large Language Models for Code, LLM4Code 2024*, 94-101. <https://doi.org/10.1145/3643795.3648389>
- Doughty, J., Wan, Z., Bompelli, A., Qayum, J., Wang, T., Zhang, J., Zheng, Y., Doyle, A., Sridhar, P., Agarwal, A., Bogart, C., Keylor, E., Kultur, C., Savelka, J., y Sakr, M. (2024). A Comparative Study of AI-Generated (GPT-4) and Human-crafted MCQs in Programming Education. *ACM International Conference Proceeding Series*, 114-123. <https://doi.org/10.1145/3636243.3636256>
- Dunder, N., Lundborg, S., Wong, J., y Viberg, O. (2024). Kattis vs ChatGPT: Assessment and evaluation of programming tasks in the age of artificial intelligence. *ACM International Conference Proceeding Series*, 821-827. <https://doi.org/10.1145/3636555.3636882>
- Espina-Romero Romero, L. C. (2022). Procesos de Enseñanza-Aprendizaje Virtual durante la COVID-19: Una revisión bibliométrica. *Revista de Ciencias Sociales (Ve)*, XXVIII(3), 345-361. <https://doi.org/10.31876/rcs.v28i3.38479>
- Frankford, E., Sauerwein, C., Bassner, P., Krusche, S., y Breu, R. (2024). AI-Tutoring in software engineering education experiences with large language models in programming assessments. *Proceedings - International Conference on Software Engineering*, 309-319. <https://doi.org/10.1145/3639474.3640061>
- Galindo, A. (2023). Integración de la inteligencia artificial en la enseñanza de las artes plásticas. *Revista de Ciencias Sociales (Ve)*, XXIX(4), 17-29. <https://doi.org/10.31876/rcs.v29i4.41256>
- Getenet, S. (2024). Pre-service teachers and ChatGPT in multistrategy problem-solving: Implications for mathematics teaching in primary schools. *International Electronic Journal of Mathematics Education*, 19(1), em0766. <https://doi.org/10.29333/iejme/14141>
- González, G. A., Martínez, L. E., Muegues, W. S., y Verdecía, L. J. (2025). Inteligencia Artificial y su impacto sobre la gerencia estratégica y la cultura investigativa. *Revista de Ciencias Sociales (Ve)*, XXXI(E-11), 312-322. <https://doi.org/10.31876/rcs.v31i.44003>
- Gouia-Zarrad, R., y Gunn, C. (2024). Enhancing students' learning experience in mathematics class through ChatGPT. *International Electronic Journal of Mathematics Education*, 19(3), em0781. <https://doi.org/10.29333/iejme/14614>
- Haindl, P., y Weinberger, G. (2024). Students' experiences of using ChatGPT in an undergraduate programming course. *IEEE Access*, 12, 43519-43529. <https://doi.org/10.1109/ACCESS.2024.3380909>
- He, Z., Nguyen, T., Miari, T., Aliasgari, M., Rafatirad, S., y Sayadi, H. (2024). The AI Companion in Education: Analyzing the Pedagogical Potential of ChatGPT in Computer Science and Engineering. *IEEE Global Engineering Education Conference, EDUCON, Kos Island, Greece*, 1-10. <https://doi.org/10.1109/EDUCON60312.2024.10578820>
- Hou, I., Mettille, S., Man, O., Li, Z., Zastudil, C., y MacNeil, S. (2024). The effects of Generative AI on computing

- students' help-seeking preferences. *ACM International Conference Proceeding Series*, 39-48. <https://doi.org/10.1145/3636243.3636248>
- Hou, X., Wu, Z., Wang, X., y Ericson, B. J. (2024). CodeTailor: LLM-Powered Personalized Parsons Puzzles for Engaging Support While Learning Programming. *L@S2024-Proceedings of the 11th ACM Conference on Learning @ Scale*, 51-62. <https://doi.org/10.1145/3657604.3662032>
- Humble, N., Boustedt, J., Holmgren, H., Milutinovic, G., Seipel, S., y Östberg, A.-S. (2023). Cheaters or AI-Enhanced Learners: Consequences of ChatGPT for Programming Education. *Electronic Journal of E-Learning*, 22(2), 16-29. <https://doi.org/10.34190/ejel.21.5.3154>
- Husain, A. J. A. (2024). Potentials of ChatGPT in computer programming: insights from programming instructors. *Journal of Information Technology Education: Research*, 23, 2. <https://doi.org/10.28945/5240>
- Jacobs, S., y Jaschke, S. (2024). Evaluating the Application of Large Language Models to Generate Feedback in Programming Education. *IEEE Global Engineering Education Conference, EDUCON*, Kos Island, Greece, 1-5. <https://doi.org/10.1109/EDUCON60312.2024.10578838>
- Jin, J., y Kim, M. (2023). GPT-Empowered Personalized eLearning System for Programming Languages. *Applied Sciences (Switzerland)*, 13(23), 12773. <https://doi.org/10.3390/app132312773>
- Jordan, M., Ly, K., y Raj, A. G. S. (2024). Need a Programming Exercise Generated in Your Native Language? ChatGPT's Got Your Back: Automatic Generation of Non-English Programming Exercises Using OpenAI GPT-3.5. *SIGCSE 2024 - Proceedings of the 55th ACM Technical Symposium on Computer Science Education*, 1, 618-624. <https://doi.org/10.1145/3626252.3630897>
- Joshi, I., Budhiraja, R., Dev, H., Kadia, J., Atallah, M. O., Mitra, S., Akolekar, H. D., y Kumar, D. (2024). ChatGPT in the classroom: An analysis of its strengths and weaknesses for solving undergraduate computer science questions. *SIGCSE 2024 - Proceedings of the 55th ACM Technical Symposium on Computer Science Education*, 1, 625-631. <https://doi.org/10.1145/3626252.3630803>
- Jošt, G., Taneski, V., y Karakatič, S. (2024). The Impact of Large Language Models on Programming Education and Student Learning Outcomes. *Applied Sciences*, 14(10), 4115. <https://doi.org/10.3390/app14104115>
- Karjanto, N. (2023). Investigating difficulties and enhancing understanding in linear algebra: Leveraging SageMath and ChatGPT for (orthogonal) diagonalization and singular value decomposition. *Mathematical Biosciences and Engineering*, 20(9), 16551-16595. <https://doi.org/10.3934/mbe.2023738>
- Kazemitabaar, M., Ye, R., Wang, X., Henley, A. Z., Denny, P., Craig, M., y Grossman, T. (May 11, 2024). CodeAid: Evaluating a Classroom Deployment of an LLM-based Programming Assistant that Balances Student and Educator Needs. *Conference on Human Factors in Computing Systems - Proceedings*, 650. <https://doi.org/10.1145/3613904.3642773>
- Kimmel, B., L., A., Yaro, L., Gipson, B., Taylor, R., Osa-Asante, S. K., Vaught, H., Wininger, G., y Yamaguchi, C. (May 11, 2024). Enhancing Programming Error Messages in Real Time with Generative AI. *Conference on Human Factors in Computing Systems*

- *Proceedings*, 608. <https://doi.org/10.1145/3613905.3647967>
- Kosar, T., Ostojić, D., Liu, Y. D., y Mernik, M. (2024). Computer science education in ChatGPT era: Experiences from an experiment in a programming course for novice programmers. *Mathematics*, 12(5), 629. <https://doi.org/10.3390/math12050629>
- Lee, J. Y. (2023). Can an artificial intelligence chatbot be the author of a scholarly article? *Science Editing*, 10(1), 7-12. <https://doi.org/10.6087/kcse.292>
- Liao, J., Zhong, L., Zhe, L., Xu, H., Liu, M., y Xie, T. (2024). Scaffolding computational thinking with ChatGPT. *IEEE Transactions on Learning Technologies*, 17, 1668-1682. <https://doi.org/10.1109/TLT.2024.3392896>
- Liffiton, M., Sheese, B. E., Savelka, J., y Denny, P. (November 13, 2023). CodeHelp: Using large language models with guardrails for scalable support in programming classes. *ACM International Conference Proceeding Series*, 8. <https://doi.org/10.1145/3631802.3631830>
- Logacheva, E., Hellas, A., Prather, J., Sarsa, S., y Leinonen, J. (2024). Evaluating Contextually Personalized Programming Exercises Created with Generative AI. *ICER 2024 - ACM Conference on International Computing Education Research*, 1, 95-113. <https://doi.org/10.1145/3632620.3671103>
- López-Fernández, D., y Vergaz, R. (2024). Adoption and impact of ChatGPT in computer science education: A case study on a database administration course. *AI*, 5(4), 2321-2337. <https://doi.org/10.3390/ai5040114>
- MacNeil, S., Denny, P., Tran, A., Leinonen, J., Bernstein, S., Hellas, A., Sarsa, S., y Kim, J. (2024). Decoding Logic Errors: A Comparative Study on Bug Detection by Students and Large Language Models. *ACM International Conference Proceeding Series*, 11-18. <https://doi.org/10.1145/3636243.3636245>
- Marín-González Cristina, F., Cuba-Romero, G. C., y Larios-Ariza, M. A. (2025). Artificial intelligence applied to the management of dyslexia cases in university settings: A systematic review. *Revista de Ciencias Sociales (Ve)*, XXXI(3), 38-56. <https://doi.org/10.31876/rcs.v31i3.44268>
- Meissner, R., Pögel, A., Ihsberner, K., Grützmüller, M., Tornack, S., Thor, A., Pengel, N., Wollersheim, H.-W., y Hardt, W. (2024). LLM-generated competence-based e-assessment items for higher education mathematics: methodology and evaluation. *Frontiers in Education*, 9, 1427502. <https://doi.org/10.3389/educ.2024.1427502>
- Mendonça, N. C. (2024). Evaluating ChatGPT-4 Vision on Brazil's National Undergraduate Computer Science Exam. *ACM Transaction on Computing Education*, 24(3), 37. <https://doi.org/10.1145/3674149>
- Montella, R., De Vita, C. G., Mellone, G., Ciricillo, T., Caramiello, D., Di Luccio, D., Kosta, S., Damaševičius, R., Maskeliūnas, R., Queirós, R., y Swacha, J. (2024). Leveraging large language models to support authoring gamified programming exercises. *Applied Sciences*, 14(18), 8344. <https://doi.org/10.3390/app14188344>
- Morales, G. P., Quispe, A., Guía, T., y Quispe, S. M. T. (2025). Realidad Aumentada y aprendizaje interactivo en estudiantes universitarios de informática. *Revista de Ciencias Sociales (Ve)*, XXXI(E-11), 436-444. <https://doi.org/10.31876/rcs.v31i.44011>
- Mustafa, M. Y., Tlili, A., Lampropoulos,

- G., Huang, R., Jandrić, P., Zhao, J., Salha, S., Xu, L., Panda, S., Kinshuk, López-Pernas, S., y Saqr, M. (2024). A systematic review of literature reviews on artificial intelligence in education (AIED): a roadmap to a future research agenda. *Smart Learning Environments*, 11, 59. <https://doi.org/10.1186/s40561-024-00350-5>
- Neumann, A. T., Yin, Y., Sowe, S., Decker, S., y Jarke, M. (2025). An LLM-Driven Chatbot in Higher Education for databases and information systems. *IEEE Transactions on Education*, 68(1), 103-116. <https://doi.org/10.1109/TE.2024.3467912>
- Ouh, E. L., Gan, B. K. S., Shim, K. J., y Wlodkowski, S. (2023). ChatGPT, can you generate solutions for my coding exercises? An evaluation on its effectiveness in an undergraduate java programming course. *Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE*, 1, 54-60. <https://doi.org/10.1145/3587102.3588794>
- Pabreja, K., y Pabreja, N. (2024). Understanding college students' satisfaction with ChatGPT: An exploratory and predictive machine learning approach using feature engineering. *MIER Journal of Educational Studies Trends & Practices*, 14(1), 37-63. <https://doi.org/10.52634/mier/2024/v14/i1/2568>
- Padiyath, A., Hou, X., Pang, A., Viramontes, D., Gu, X., Nelson-Fromm, T., Wu, Z., Guzdial, M., y Ericson, B. (2024). Insights from Social Shaping Theory: The Appropriation of Large Language Models in an Undergraduate Programming Course. *Proceedings of the 2024 ACM Conference on International Computing Education Research*, 1, 114-130. <https://doi.org/10.1145/3632620.3671098>
- Pankiewicz, M., y Baker, R. S. (2024). Navigating Compiler Errors with AI Assistance-A Study of GPT Hints in an Introductory Programming Course. *ITiCSE 2024: Proceedings of the 2024 on Innovation and Technology in Computer Science Education*, 1, 94-100. <https://doi.org/10.1145/3649217.3653608>
- Perezchica-Vega, J. E., Sepúlveda-Rodríguez, J. A., y Román-Méndez, A. D. (2024). Generative artificial intelligence in higher education: uses and opinions of teachers. *European Public & Social Innovation Review*, 9, 1-20. <https://doi.org/10.31637/epsir-2024-593>
- Phung, T., Pădurean, V.-A., Singh, A., Brooks, C., Cambronero, J., Gulwani, S., Singla, A., y Soares, G. (2024). Automating Human Tutor-Style Programming Feedback: Leveraging GPT-4 tutor model for hint generation and GPT-3.5 student model for hint validation. *ACM International Conference Proceeding Series*, 12-23. <https://doi.org/10.1145/3636555.3636846>
- Prather, J., Denny, P., Leinonen, J., Becker, B. A., Albluwi, I., Craig, M., Keuning, H., Kiesler, N., Kohn, T., Luxton-Reilly, A., MacNeil, S., Petersen, A., Pettit, R., Reeves, B. N., y Savelka, J. (2023). The Robots are Here: Navigating the Generative AI Revolution in Computing Education. *ITiCSE-WGR 2023 - Proceedings of the 2023 Working Group Reports on Innovation and Technology in Computer Science Education*, 108-159. <https://doi.org/10.1145/3623762.3633499>
- Quiñones, A. J., y Huiman, H. E. (2022). Resolución de problemas con el método matemático de Polya: La aventura de aprender. *Revista de Ciencias Sociales (Ve)*, XXVIII(E-5), 75-86. <https://doi.org/10.31876/rcs.v28i.38146>
- Rajala, J., Hukkanen, J., Hartikainen, M., y

- Niemelä, P. (2023). “Call me Kiran” ChatGPT as a tutoring Chatbot in a computer science course. *ACM International Conference Proceeding Series: Mindtrek '23: Proceedings of the 26th International Academic Mindtrek Conferenc*, 83-94. <https://doi.org/10.1145/3616961.3616974>
- Rasnayaka, S., Wang, G., Shariffdeen, R., e Iyer, G. N. (2024). An Empirical study on usage and perceptions of LLMs in a software engineering project. *Proceedings - 2024 International Workshop on Large Language Models for Code, LLM4Code 2024*, 111-118. <https://doi.org/10.1145/3643795.3648379>
- Richards, M., Waugh, K., Slaymaker, M., Petre, M., Woodthorpe, J., y Gooch, D. (2024). Bob or Bot: Exploring ChatGPT’s Answers to University Computer Science Assessment. *ACM Transactions on Computing Education*, 24(1), 5. <https://doi.org/10.1145/3633287>
- Robleda, G. (2019). How to analyze and write the results of a systematic review. *Enfermería Intensiva*, 30(4), 192-195. <https://doi.org/10.1016/j.enfi.2019.09.001>
- Roest, L., Keuning, H., y Jeurig, J. (2023). *Next-Step Hint Generation for Introductory Programming Using Large Language Models*. arXiv:2312.10055. <https://doi.org/10.48550/arXiv.2312.10055>
- Sánchez-Ruiz, J., y Larrea-Silva, J. (2022). Factores que determinan la satisfacción estudiantil en tiempos de pandemia Covid-19: Revisión sistemática. *Revista de Ciencias Sociales (Ve)*, XXVIII(E-6), 32-44. <https://doi.org/10.31876/rcs.v28i.38817>
- Sánchez-Ruiz, L. M., Moll-López, S., Nuñez-Pérez, A., Moraño-Fernández, J. A., y Vega-Fleitas, E. (2023). ChatGPT challenges blended learning methodologies in engineering education: A case study in mathematics. *Applied Sciences*, 13(10), 6039. <https://doi.org/10.3390/app13106039>
- Sarsa, S., Denny, P., Hellas, A., y Leinonen, J. (2022). Automatic generation of programming exercises and code explanations using large language models. *ICER 2022 - Proceedings of the 2022 ACM Conference on International Computing Education Research*, 1, 27-43. <https://doi.org/10.1145/3501385.3543957>
- Savelka, J., Agarwal, A., An, M., Bogart, C., y Sakr, M. (2023). Thrilled by Your Progress! Large Language Models (GPT-4) No Longer Struggle to Pass Assessments in Higher Education Programming Courses. *ICER 2023 - Proceedings of the 2023 ACM Conference on International Computing Education Research*, 1, 78-92. <https://doi.org/10.1145/3568813.3600142>
- Savelka, J., Ashley, K. D., Gray, M. A., Westermann, H., y Xu, H. (June 2023). Can GPT-4 support analysis of textual data in tasks requiring highly specialized domain expertise? *Proceedings of the Sixth Workshop on Automated Semantic Analysis of Information in Legal Text (ASAIL)*, 23, Braga, Portugal, 1-12. <https://ceur-ws.org/Vol-3441/paper1.pdf>
- Segal, R., y Biton, Y. (2024). The contribution that utilizing Generative AI for problem posing makes to pre-service high school mathematics teachers’ TPACK. *International Journal of Education in Mathematics, Science and Technology*, 12(6), 1559-1582. <https://doi.org/10.46328/ijemst.4591>
- Shen, Y., Ai, X., Soosai, A. G., Leo, R. J., y Syamkumar, M. (2024).

- Implications of ChatGPT for Data Science Education. *SIGCSE 2024 - Proceedings of the 55th ACM Technical Symposium on Computer Science Education, 1*, 1230-1236. <https://doi.org/10.1145/3626252.3630874>
- Singh, H., Tayarani-Najaran, M.-H., y Yaqoob, M. (2023). Exploring computer science students' perception of ChatGPT in Higher Education: A descriptive and correlation study. *Education Sciences, 13*(9), 924. <https://doi.org/10.3390/educsci13090924>
- Sterne, J. A. C., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., Cates, C. J., Cheng, H.-Y., Corbett, M. S., Eldridge, S. M., Emberson, J. R., Hernán, M. A., Hopewell, S., Hróbjartsson, A., Junqueira, D. R., Jüni, P., Kirkham, J. J., Lasserson, T., Li, T., ... Higgins, J. P. T. (2019). RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ, 366*, 14898. <https://doi.org/10.1136/bmj.14898>
- Sun, D., Boudouaia, A., Zhu, C., y 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*, 14. <https://doi.org/10.1186/s41239-024-00446-5>
- Udias, A., Alonso-Ayuso, A., Alfaro, C., Algar, M. J., Cuesta, M., Fernández-Isabel, A., Gómez, J., Lanco, C., Cano, E. L., Martín, I., y Ortega, F. (2024). ChatGPT's performance in university admissions tests in mathematics. *International Electronic Journal of Mathematics Education, 19*(4), em0795. <https://doi.org/10.29333/iejme/15517>
- Wahba, F., Ajlouni, A. O., y Abumosa, M. A. (2024). The impact of ChatGPT-based learning statistics on undergraduates' statistical reasoning and attitudes toward statistics. *Eurasia Journal of Mathematics, Science and Technology Education, 20*(7), em2468. <https://doi.org/10.29333/ejmste/14726>
- Wilson, D. B. (2019). Systematic coding for research synthesis. In H. Cooper, L. V. Hedges y J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (3 ed., pp. 154-174). Russell Sage Foundation.
- Xue, Y., Chen, H., Bai, G. R., Tairas, R., y Huang, Y. (2024). Does ChatGPT help with introductory programming? An experiment of students using ChatGPT in CS1. *Proceedings - International Conference on Software Engineering, 331-341*. <https://doi.org/10.1145/3639474.3640076>
- Yi, L., Liu, D., Jiang, T., y Xian, Y. (2025). The Effectiveness of AI on K-12 Students' Mathematics Learning: A Systematic Review and Meta-Analysis. *International Journal of Science and Mathematics Education, 23*(4), 1105-1126. <https://doi.org/10.1007/s10763-024-10499-7>
- Zastudil, C., Rogalska, M., Kapp, C., Vaughn, J., y MacNeil, S. (2023). Generative AI in Computing Education: Perspectives of Students and Instructors. *arXiv:2308.04309*. <https://doi.org/10.48550/arXiv.2308.04309>