



Impacto de la IA en la enseñanza de la programación: Una revisión sistemática

Miguel-Leonardo Avila-Avila¹
Edison Ferney Gutierrez Buitrago²

Resumen

Se analizó el impacto del uso de herramientas de inteligencia artificial en la enseñanza de la programación en programas de Ingeniería de Sistemas y disciplinas relacionadas. Se utilizó la metodología PRISMA para seleccionar los estudios a analizar. En estos, se observó que las herramientas IA más utilizadas fueron ChatGPT y Copilot, las cuales han demostrado, mayoritariamente, impactos positivos en los estudios para potenciar la educación. Así mismo, estas herramientas permiten aumentar la eficiencia en la resolución de problemas, la motivación y confianza, debido a que ofrecen retroalimentación inmediata y asistencia personalizada a los estudiantes. Adicionalmente, contribuyen con el desarrollo de habilidades cognitivas de orden superior y permiten una reducción en la carga cognitiva. Sin embargo, debe considerarse que un uso no regulado o sin acompañamiento pedagógico puede generar impactos negativos, los cuales se encaminan a generar dependencia excesiva, realizar plagio, la pérdida del pensamiento crítico y el riesgo a la generación de información falsa.

Palabras clave: educación; inteligencia artificial; programación.

Impact of Artificial Intelligence on Programming Education: A Systematic Review

Abstract

The impact of using artificial intelligence tools on the teaching of programming in Systems Engineering programs and related disciplines was analyzed. The PRISMA methodology was used to select the studies to be reviewed. Among these, the most used AI tools were ChatGPT and Copilot, which have mostly shown positive impacts on education. These tools

¹ Universidad Pedagógica y Tecnológica de Colombia, Tunja, Boyacá. miguel.avila04@uptc.edu.co

² Universidad Pedagógica y Tecnológica de Colombia, Tunja, Boyacá. edison.gutierrez01@uptc.edu.co

help increase efficiency in problem-solving, motivation, and confidence by offering immediate feedback and personalized assistance to students. Additionally, they contribute to the development of higher-order cognitive skills and help reduce cognitive load. However, it should be considered that unregulated use or use without pedagogical guidance can generate negative impacts, such as excessive dependence, plagiarism, loss of critical thinking, and the risk of generating false information.

Keywords: artificial intelligence; education; programming.

Introducción

En los últimos años, ha sido evidente que la incorporación de herramientas de Inteligencia Artificial (IA) en diferentes aspectos del diario vivir, como en la educación, generando un cambio significativo en la forma de enseñanza y aprendizaje. En programas relacionados con Ingeniería de Sistemas, específicamente, en el campo de la Ingeniería de Software, los LLMs como ChatGPT, Claude y Copilot han comenzado a ser utilizados en proyectos académicos, apoyando tareas de diseño, programación y depuración (Kharrufa et al., 2024). Sin embargo, aún existe la incertidumbre sobre su impacto real en la comprensión de los conceptos del ciclo de vida del software por los estudiantes (Gaitantzi & Kazanidis, 2025). Además, se han identificado riesgos asociados al uso de Modelos de Lenguaje de Gran Escala (LLM, por las siglas en inglés de Large Language Models), como la dependencia excesiva, la disminución de las capacidades críticas y la propagación de información incorrecta (Ferino et al., 2025).

Sin la construcción de lineamientos claros para la integración de herramientas IA en el proceso de enseñanza-aprendizaje, los ciclos de formación corren el riesgo de no preparar adecuadamente a los futuros profesionales. Esto podría derivar en profesionales dependientes de las herramientas, con menor capacidad de innovación y resolución de problemas de manera autónoma (Kharrufa et al., 2024).

Por tal razón, es necesario realizar una investigación que evalúe el impacto del uso de modelos de lenguaje de gran escala (LLMs) en la enseñanza y el aprendizaje en programas de Ingeniería de Sistemas y disciplinas relacionadas, con el fin de identificar beneficios, riesgos y lineamientos que permitan su uso responsable y efectivo en la educación.

Metodología

Al tratarse de una revisión sistemática, se hizo uso de la metodología PRISMA (Page et al., 2021), debido a que es ampliamente utilizada en diferentes campos de estudio (PRISMA,

2020). PRISMA permite la documentación de manera estructurada en cada una de sus fases y se adapta al tipo de estudio a realizar. En la Tabla 1, se muestran los criterios que se tomaron en cuenta para realizar la selección de los artículos relacionados al tema de investigación.

Tabla 1. Criterios de selección para la búsqueda de artículos.

Nº	Criterio	Justificación
1	Artículos redactados en inglés.	Identificar artículos en inglés, porque este es el idioma predominante en la publicación científica internacional, lo que permite acceder a mayor cantidad de evidencia útil (PRISMA, 2020; PuntoEdu, 2016).
2	Publicaciones comprendidas entre enero de 2020 y septiembre de 2025.	Hallar artículos publicados entre septiembre de 2020 y septiembre de 2025, porque, en los últimos cinco años el campo de la inteligencia artificial ha tenido un desarrollo significativo (Vizcaino, 2020), y por lo tanto las investigaciones publicadas en este rango de fechas serán más provechosas.
3	Estudios centrados en el impacto que tiene el uso de herramientas de inteligencia artificial (positivo, negativo, mixto)	Identificar estudios centrados en el impacto que tiene el uso de herramientas de inteligencia artificial, para identificar las diversas perspectivas del fenómeno estudiado.
4	Estudios que aborden la relación entre el uso de herramientas de inteligencia artificial y el desarrollo (programación de software).	Identificar artículos enfocados en el impacto que tiene el uso de herramientas de inteligencia artificial en el aprendizaje de desarrollo y programación de software porque es un objetivo principal de la investigación.
5	Estudios que aborden la relación entre el uso de herramientas de inteligencia artificial y el aprendizaje de estudiantes de educación superior (ingeniería de Sistemas y disciplinas afines).	Este grupo de estudiantes fue seleccionado debido a que habitualmente usan algún lenguaje de programación y hacen uso de herramientas de inteligencia artificial como apoyo en su proceso de aprendizaje.
6	Investigaciones relacionadas únicamente al ámbito educativo, no en el ambiente laboral.	Omitir artículos enfocados en el área laboral/industrial, porque el interés de la investigación está en el impacto pedagógico de la IA. Esta delimitación asegura coherencia con el objetivo de la revisión.
7	Artículos de revisiones sistemáticas o investigaciones empíricas	Enfocar la investigación en artículos de revisiones sistemáticas o investigaciones empíricas. Las revisiones sistemáticas aportan síntesis de evidencia existente y del estado del estudio (Global, s.f.), mientras que los estudios empíricos permiten observar resultados directos y medibles (Atlas, 2025; Velasquez, 2023). Al

Impacto de la IA en la enseñanza de la programación: Una revisión sistemática

Nº	Criterio	Justificación
		restringirse a estos tipos de estudios, se asegura validez científica y solidez en los hallazgos de la revisión.

Para la correcta selección de los artículos a incluir en el análisis sistemático, se realizó un filtro de los artículos, utilizando diversas estrategias de búsqueda, con fin de encontrar los artículos con mayor relación a los criterios de estudio (Tabla 2). Para ello se realizaron búsquedas en bases de datos tales como OpenAlex, Springer Nature y Scopus.

Tabla 2. Estrategia de búsqueda para la base de datos seleccionada.

Base de datos	Estrategia de búsqueda empleada
SCOPUS	TITLE-ABS-KEY(("AI" OR "artificial intelligence") AND (university OR "k-12" OR "primary school" OR "baccalaureate" OR "higher education" OR bachelor OR "system engineering" OR "computer science" OR "software engineering") AND ("education" OR learning) AND (impact) AND ("programming" OR "software development" OR "program code" OR "frontend development" OR "backend development" OR "software architecture" OR software) AND ("student") AND ("systematic review" OR "literature review" OR "scoping review" OR "meta-analysis" OR "survey of research" OR "state of the art" OR "empirical study" OR "case study" OR "experimental study" OR "quasi-experimental" OR "mixed methods" OR "qualitative study" OR "quantitative study" OR "field study") AND NOT ("job project" OR "job" OR "profesional work" OR "job industry" OR "industry") AND NOT(database OR "data design" OR "data model" OR "data architecture") AND NOT("requirement engineering" OR requirement) AND NOT("design learning" OR "diagram learning" OR "diagram" OR "model learning")) AND PUBYEAR > 2019 AND PUBYEAR < 2026
SPRINGER NATURE	("AI" OR "artificial intelligence") AND (university OR "k-12" OR "primary school" OR "baccalaureate" OR "higher education" OR bachelor OR "system engineering" OR "computer science") AND ("education" OR learning) AND ("student") AND (impact) AND ("programming" OR "software development" OR "program code" OR "frontend development" OR "backend development" OR "software architecture" OR software) AND ("systematic review" OR "literature review" OR "scoping review" OR "meta-analysis" OR "survey of research" OR "state of the art" OR "empirical study" OR "case study" OR "experimental study" OR "quasi-experimental" OR "mixed methods" OR "qualitative study" OR "quantitative study" OR "field study") AND NOT ("job project" OR "job" OR "profesional work" OR "business" OR "job industry" OR "industry") AND NOT(database OR "data design" OR "data model" OR "data architecture") AND NOT("requirement engineering" OR requirement) AND NOT ("design learning" OR "diagram learning" OR diagram OR "model learning")

Base de datos	Estrategia de búsqueda empleada
OPENALEX	("AI" OR "artificial intelligence") AND (university OR "k-12" OR "primary school" OR "baccalaureate" OR "higher education" OR bachelor OR "system engineering" OR "computer science") AND ("education" OR learning) AND ("student") AND (impact) AND ("programming" OR "software development" OR "program code" OR "frontend development" OR "backend development" OR "software architecture" OR software) AND ("systematic review" OR "literature review" OR "scoping review" OR "meta-analysis" OR "survey of research" OR "state of the art" OR "empirical study" OR "case study" OR "experimental study" OR "quasi-experimental" OR "mixed methods" OR "qualitative study" OR "quantitative study" OR "field study") AND NOT ("job project" OR "job" OR "profesional work" OR "business" OR "job industry" OR "industry") AND NOT(database OR "data design" OR "data model" OR "data architecture") AND NOT("requirement engineering" OR requirement) AND NOT ("design learning" OR "diagram learning" OR diagram OR "model learning")

Al realizar las consultas en las bases de datos anteriormente mencionadas y utilizando las estrategias de búsqueda planteadas, en la Tabla 3 se encontraron 97 artículos en Scopus, 268 en Springer y 267 en OpenAlex. Sin embargo, se determinó utilizar los resultados obtenidos de Scopus debido a que los artículos encontrados en esta base de datos tuvieron mayor relación con los criterios de selección. En la inclusión de artículos se siguieron los pasos especificados en la metodología PRISMA (Page et al., 2021), divididos en las 4 fases:

- **Identificación de artículos:** Fueron identificados 97 artículos al realizar la búsqueda usando el algoritmo definido previamente, consultados el 26 de septiembre de 2025.
- **Proyección de artículos:** Luego de realizar una revisión manual no se detectó ningún artículo duplicado entre los artículos identificados. Luego, tras leer el título, 40 de los artículos fueron descartados, porque no tuvieron relación con las palabras clave de la estrategia de búsqueda ingresada. La revisión fue realizada de manera independiente por dos autores, donde cada uno revisó la mitad de los artículos y ambos se apoyaron para resolver dudas con respecto a la selección de artículos.
- **Selección de artículos:** Posteriormente, se excluyeron 39 artículos: 5 porque no se tuvo acceso a su versión completa, 4 porque pudo accederse únicamente a las versiones preliminares sin revisión de pares. Luego fueron descartados 23 artículos porque tras leer el Abstract no tenían relación con los criterios de selección. Finalmente, se excluyeron 7 artículos debido a que no cumplían con la totalidad de los criterios de selección establecidos.

- **Inclusión de artículos:** Finalmente, fueron incluidos 18 artículos para el análisis sistemático y la síntesis de resultados.

La Figura 1 muestra la aplicación de la metodología PRISMA fase por fase en la selección de artículos.

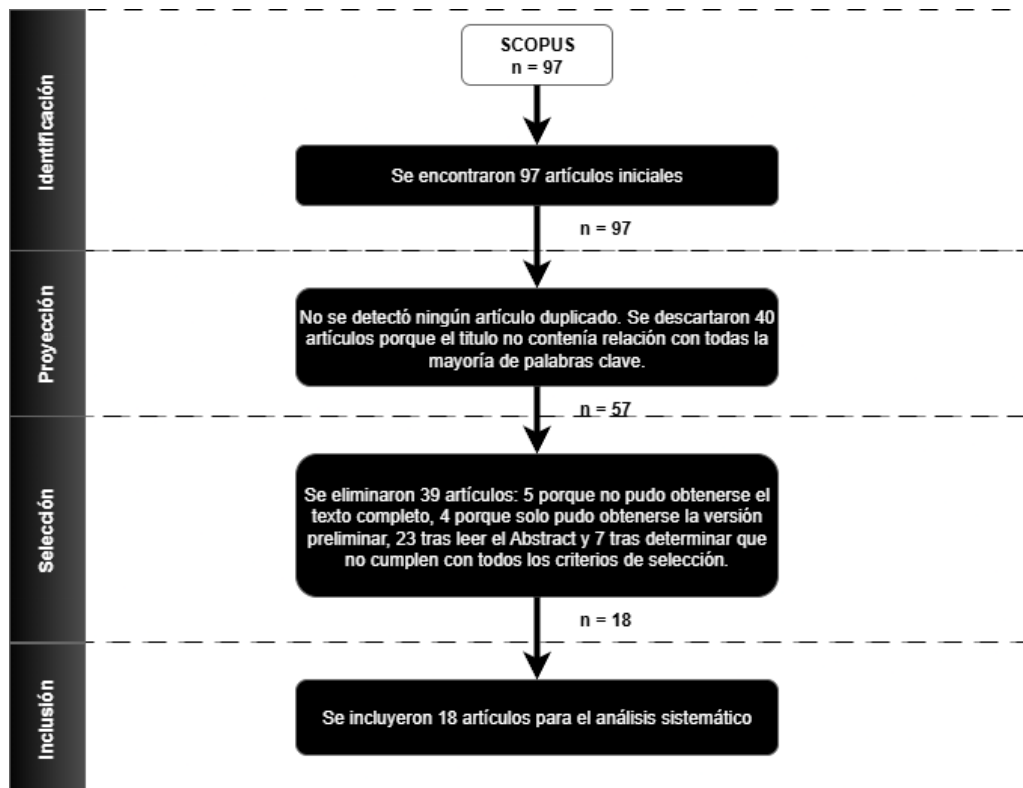


Figura 1. Procesos de identificación, proyección, selección e inclusión siguiendo la metodología PRISMA.

Resultados

En esta sección se presenta el análisis de los artículos incluidos en el estudio. Los resultados detallados a continuación incluyen aspectos como las características generales de los artículos, metodologías incluidas, herramientas utilizadas, impacto en la enseñanza y las métricas de evaluación. El objetivo es presentar una contextualización de la muestra de artículos.

Los artículos incluidos fueron publicados en el periodo entre septiembre de 2020 y septiembre de 2025, aclarando que en los años 2020 a 2022 no se publicó ningún artículo relacionado al tema de estudio, esto puede deberse a que en estos años la inteligencia artificial aún estaba en sus fases iniciales de popularidad mediática, por lo que los artículos

que midieron su impacto en la educación fueron mínimos. En la Figura 2, se observa la distribución del país de origen de los autores en el mapa mundial, donde entre más verde oscuro sea el país mayor cantidad de autores tuvo.

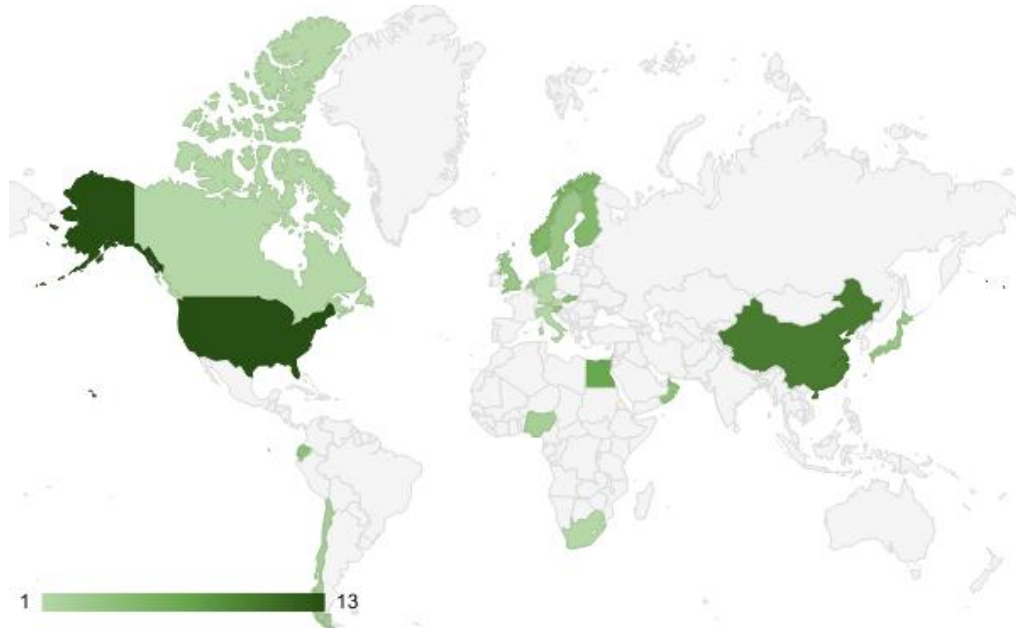


Figura 2. Distribución del país de origen de los autores.

Los artículos provienen de una gran variedad de países y de regiones como puede observarse. Muchos de ellos fueron escritos por autores de diferentes nacionalidades. La mayoría de los autores provienen de países del hemisferio norte, países que en su mayoría son líderes en tecnología, educación superior e investigación, lo que sugiere una mayor capacidad para publicar estudios sobre temas de inteligencia artificial.

Se observa que Asia lidera la producción de investigaciones, esto puede deberse a un fuerte enfoque en la educación, el cual es apoyado por inversiones significativas en tecnología, lo que facilita la rápida innovación e investigación (Chang et al., 2024).

Por otro lado, en cuanto a los países latinoamericanos, se observa que únicamente 6 autores provienen de esta región, esto se relaciona con el estado del campo de la inteligencia artificial en la región, el cual se encuentra en sus etapas iniciales y enfrenta diversos desafíos. Los estudiantes de esta región pueden tener diferentes niveles de alfabetización digital con respecto a países de Europa, Asia y Estados Unidos, lo que influye en la adopción de tecnologías de IA en la educación de la programación (Mellado & Cubillos, 2025).

Metodologías utilizadas en los estudios

En cuanto al tipo de diseño de investigación, los artículos abarcan una gran variedad de metodologías. Los diseños de investigación de los artículos pueden clasificarse así:

- Estudio empírico: Fue elegido en algunos de los artículos con el fin de evaluar de manera rigurosa el impacto cognitivo y afectivo de las herramientas de IA en un entorno académico real.
- Método mixto: Utilizado para obtener una comprensión más completa y profunda de la interacción entre las herramientas de IA y el proceso de aprendizaje, mezclando estudios cuantitativos, cualitativos y triangulando resultados.
- Revisión sistemática: Se implementó para lograr una comprensión exhaustiva, sintetizada y fundamentada del uso e impacto de la IA en la educación de la programación.
- Estudio cuasi-experimental/controlado: Utilizado por algunos autores con fin de evaluar rigurosamente el impacto causal del uso de herramientas IA en la educación.
- Estudio cuantitativo: Implementado con fin de medir, comparar y cuantificar los efectos del uso de herramientas de IA.

En la Figura 3, se observa la cantidad de artículos que utilizaron cada diseño de investigación.

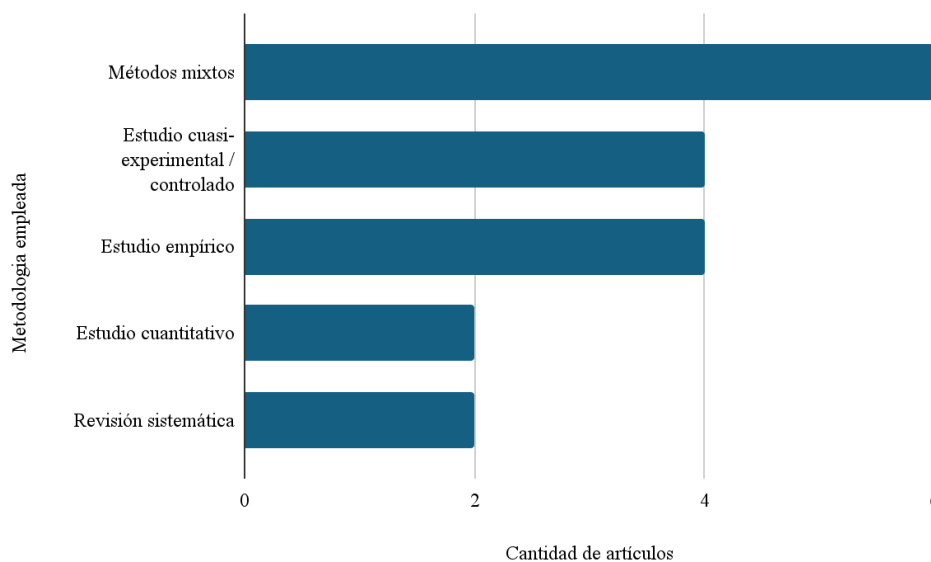


Figura 3. Cantidad de artículos clasificados por tipo de metodología.

Esta variedad de diseños de investigación refleja la amplitud de enfoques en el campo de investigación. Siendo los métodos mixtos el diseño más utilizado, lo que demuestra que la mayoría de los artículos buscan obtener una comprensión más profunda y fundamentada del impacto del fenómeno estudiado utilizando investigación cuantitativa y cualitativa (Atta et al., 2025; Abdulla et al., 2024; Nathaniel et al., 2025). Las metodologías mixtas permiten a los investigadores identificar estrategias de integración de las herramientas de IA sin comprometer la integridad en el proceso de aprendizaje, analizando tanto las puntuaciones como los comentarios y las reflexiones de los estudiantes (Abdulla et al., 2024).

Variables de relevancia en los estudios

Es fundamental analizar con precisión cuándo, cómo y para quién la integración de una herramienta IA en la enseñanza de la programación es realmente beneficiosa o perjudicial (Li et al., 2023; Wang et al. 2023; Baláž et al., 2024). Para lograr esto primero debemos analizar la herramienta de inteligencia artificial utilizada en los artículos, el lenguaje de programación de enfoque y el nivel educativo de los estudiantes sobre los cuales se realizó el estudio. En la Figura 4 se presenta la relación de las herramientas IA que se utilizaron en el estudio realizado en cada uno de los artículos seleccionados, además de los diferentes niveles educativos que se tomaron en cuenta para realizar el estudio.

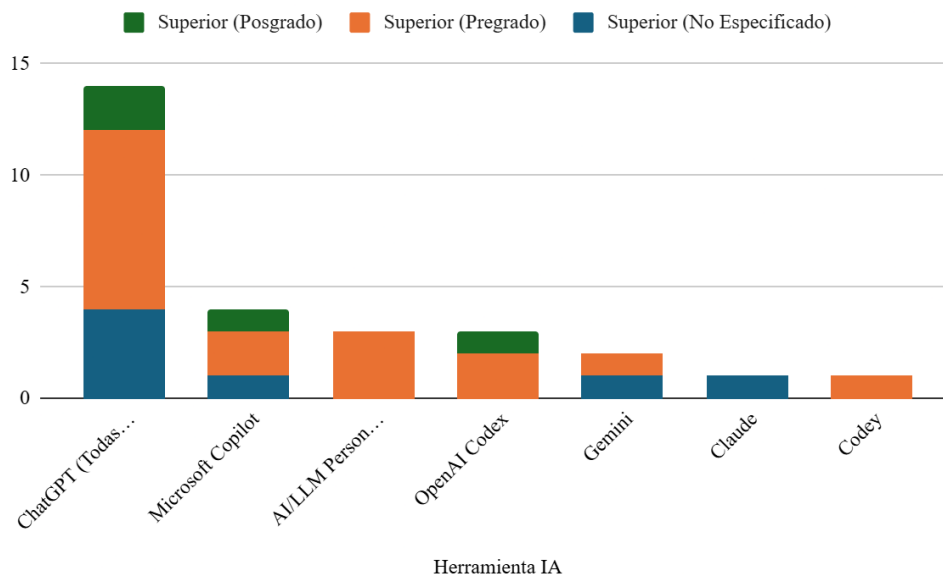


Figura 4. IA utilizada respecto al nivel educativo de enfoque

Los resultados obtenidos son los siguientes:

- La herramienta IA más utilizada por los diferentes niveles educativos fue ChatGPT (en cualquiera de sus versiones), esto se debe a la amplia accesibilidad de la herramienta, su lanzamiento oportuno y popular, y sus capacidades generales como asistente del aprendizaje (Chang et al., 2024).
- A pesar de que existen herramientas más especializadas para la programación como Copilot y OpenAI Codex (Chang et al., 2024), ChatGPT se destacó entre los estudiantes de los diferentes niveles por ser un asistente general en tareas de enseñanza e investigación (Atta et al., 2025).
- Aunque modelos como Gemini, caracterizados por la claridad en sus respuestas en niveles educativos más simples (k-12), ChatGPT superó a dichas herramientas en profundidad y precisión en tareas técnicas (Llerena et al., 2024). Incluso en un estudio Gemini y Copilot fueron utilizados únicamente como alternativa cuando ChatGPT no estuvo disponible (Nathaniel et al., 2025).
- Las herramientas como Claude y Codey, fueron las menos utilizadas por los distintos niveles educativos debido a que son modelos recientes por lo que muchas investigaciones fueron publicadas antes de los lanzamientos de sus últimas versiones (Oyelere & Aruleba, 2025; Chang et al., 2024).
- Se observa que en varias de las investigaciones los investigadores tanto de pregrado como de posgrado trabajaron de forma paralela con más de una herramienta de IA, debido a la funcionalidad especializada de cada una y la necesidad de una evaluación comparativa entre ellas (Abdulla et al., 2024).

Por otra parte, en la Figura 5 se presenta un mapa de calor que muestra la relación entre las herramientas de inteligencia artificial utilizadas y los diferentes lenguajes de programación empleados en los artículos analizados.

ChatGPT	3	3	1	1	1
Copilot	2	0	0	1	0
AI/LLM Personalizado	1	0	1	1	0
Gemini	2	0	0	0	0
OpenAI Codex	0	1	0	1	0
Claude	1	0	0	0	0
Codey	0	1	0	0	0
	Python	Java	C#	C++	PHP

Figura 5. IA utilizada según los lenguajes de programación

Los resultados obtenidos son los siguientes:

- Los lenguajes de programación a los cuales se asoció una mayor cantidad de herramientas de inteligencia artificial fueron python y java, debido a que son los más utilizados en la enseñanza de programación (Li et al., 2023; Kosar et al., 2024; Abdulla et al., 2024).
- Múltiples de los estudios se centraron en cursos de Python, incluyendo un estudio de caso de ocho semanas con estudiantes de primer año en un curso de programación en Python. Los modelos de Chat GPT fueron comparados con tutores humanos utilizando problemas de programación en Python y programas con errores reales (Ma et al., 2024).
- Un estudio de campo de un semestre utilizó una herramienta basada en LLM (CodeTutor) configurada para actuar como asistente de enseñanza en una clase de Python de nivel introductorio (Lyu et al., 2024).
- En el caso de Java, ChatGPT y otros LLMs fueron evaluados para generar soluciones en un curso de programación de nivel universitario (Mellado & Cubillos, 2025). Además, se evaluó el impacto de la IA en la enseñanza de pruebas de software utilizando el juego Code Defenders (Mezzaro et al., 2024).
- Por otro lado, PHP fue el lenguaje de programación menos asociado a herramientas de inteligencia artificial, debido a que, para su población de estudiantes novatos, la inclusión de IA no ofrecía una ventaja significativa en las etapas iniciales de aprendizaje (Mellado & Cubillos, 2025).

Por último, en la Figura 6 se presenta la relación entre lenguajes de programación y niveles educativos analizados.

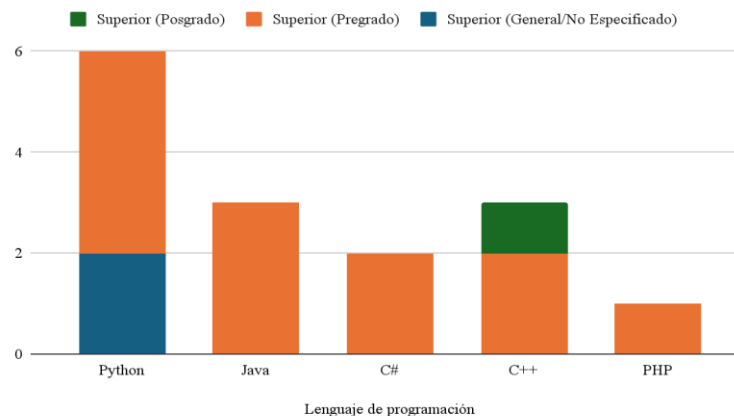


Figura 6. Lenguajes de programación según el nivel educativo

Los resultados obtenidos son los siguientes:

- El lenguaje de programación más utilizado por los estudiantes en los artículos enfocados en programas de pregrado fue Python, esto se debe a que este lenguaje se utiliza muy a menudo en los cursos de programación para principiantes debido a su sintaxis simple (Kosar et al., 2024). Además, Python permite que incluso aquellos sin conocimientos previos de programación puedan resolver problemas de código fácilmente (Ma et al., 2024).
- Los artículos enfocados en los programas de pregrado se orientan a aprender Python, Java, C#, C++ y PHP únicamente, esto se debe a que estos lenguajes se enfocan en habilidades clave de la programación por ejemplo la programación orientada a objetos, concepto esencial para el aprendizaje de los estudiantes (Ata et al., 2025). Otros lenguajes como Java y PHP se enfocan en el desarrollo web principalmente, una habilidad sumamente requerida (Baláž et al., 2024).
- El artículo enfocado en los cursos de posgrado se orientó únicamente al aprendizaje del lenguaje C++, sin embargo, no se debe a ninguna razón específica. Los cursos de posgrado se enfocan sobre todo en el aprendizaje de la ingeniería del software (Li et al., 2023; Mezzaro et al., 2024), donde el lenguaje de programación utilizado es simplemente una herramienta y el objetivo de aprendizaje son las habilidades cognitivas de orden superior y la aplicación de principios especializados de la disciplina (Kosar et al., 2024).
- El lenguaje de programación menos utilizado en los artículos es PHP, esto se debe a que no suele usarse para cursos introductorios en comparación con Python. Por tanto, PHP no tiene mucha mención en literatura orientada al aprendizaje de fundamentos de programación (Mellado & Cubillos, 2025; Chang et al., 2024).
- Los artículos se enfocaron únicamente en entornos de educación superior porque fue el nivel donde las consecuencias académicas (como plagio) se hicieron evidentes de inmediato tras el lanzamiento de los LLMs. Esto obligó a los investigadores a reaccionar rápidamente para abordar las implicaciones en la enseñanza (Wang et al., 2023; Azoulay et al., 2025; Chang et al., 2024).

Impacto

Los artículos determinaron una gran variedad de factores de impacto, sin embargo, para el análisis se sintetizaron los factores según la similitud entre unos y otros. Respecto a los

factores positivos, surgen de la capacidad de la IA para proporcionar apoyo inmediato y adaptable, permitiendo a los estudiantes resolver dudas, problemas en el código y facilitar la comprensión más profunda de una temática (si se usa correctamente) (Wang et al., 2025; Nathaniel et al., 2025). Dentro de ellos destacan:

- Asistencia en tareas, depuración y eficiencia
- Aumento de motivación
- Desarrollo de habilidades de orden superior
- Reducción de carga cognitiva

Y los negativos, son factores que surgen principalmente porque el uso de la IA permite a los estudiantes evitar el análisis y la práctica, necesarios para construir una comprensión de los fundamentos de la programación (Azoulay et al., 2025), resaltando:

- Riesgo de dependencia o sobre dependencia
- Riesgo de plagio o fraude
- Pérdida de pensamiento crítico
- Riesgo de información incorrecta o falsos positivos

Discusión

En el análisis de los artículos revisados se identificó una variedad de impactos derivados del uso de herramientas de inteligencia artificial en la enseñanza de programación a estudiantes de diferentes niveles educativos. Estos impactos se clasificaron principalmente en positivos, asociados a los beneficios que conlleva el uso de dichas herramientas en el aprendizaje y eficiencia, y negativos, relacionados a los riesgos o limitaciones que pueden surgir durante su implementación.

En la Figura 7 se presentan los diferentes tipos de impacto identificados en los artículos revisados. Se observan aquellos factores en los cuales ha sido beneficiosa la integración de herramientas de inteligencia artificial, tales como la asistencia de tareas, la motivación estudiantil, el desarrollo de habilidades de orden superior y la reducción de la carga cognitiva. Cada color representa un tipo de impacto reportado: positivo, doble (positivo y negativo), neutro o negativo.

Impacto de la IA en la enseñanza de la programación: Una revisión sistemática

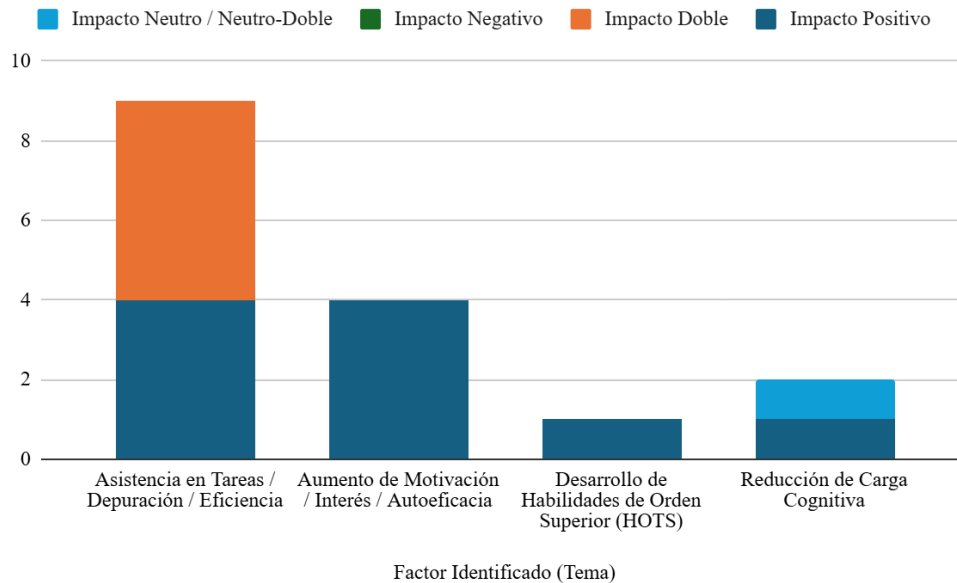


Figura 7. Impacto de los factores positivos en el aprendizaje

Los resultados obtenidos son los siguientes:

- Chat GPT y Copilot ofrecen un beneficio claro dado que ayudan a agilizar los procesos de resolución de problemas y reducir la sobrecarga cognitiva al automatizar tareas repetitivas de codificación (Azoulay et al., 2025; Oyelere & Aruleba, 2025). Dicha asistencia incluye corregir código, responder a preguntas, y proporcionar ejemplos (Ma et al., 2024).
- El aumento de la motivación fue influenciado debido a que la IA tiene la capacidad de generar respuestas claras y aclarar errores, lo que aumenta la eficiencia en la programación y la motivación en los procesos de aprendizaje (Ata et al., 2025). Para alrededor del 91% de los encuestados, el uso de la IA se ha asociado con el aumento significativo del entusiasmo por la programación (Llerena et al., 2024).
- El uso de IA tiene potencial para desarrollar y mantener las HOTS (Habilidades de orden superior), como la resolución de problemas, la creatividad y el pensamiento crítico (Nathaniel et al., 2025).
- La inteligencia artificial permite reducir la carga cognitiva al proporcionar apoyo contextualizado (Azoulay et al., 2025), y facilitar tareas de depuración y generación de código (Nathaniel et al., 2025). Además, contribuye a la asimilación de conceptos al reducir la carga cognitiva extraña (Mellado & Cubillos, 2025).

- El factor “Asistencia en tareas” fue el único que generó un impacto tanto positivo como negativo, esto se debe a que mientras las IA permiten que los estudiantes aumenten la velocidad y el rendimiento en la finalización de tareas de programación. Por otro lado, amenazan que los estudiantes comprendan conceptos profundos de la programación porque usando una herramienta IA obtienen respuestas rápidas lo que impide que primero comprendan sus acciones y el funcionamiento de una aplicación y resta valor al pensamiento crítico (Baláž et al., 2024).

En la Figura 8 se muestran los tipos de impacto negativo identificados en los artículos revisados. Se evidencian los factores en los cuales el uso de herramientas de inteligencia artificial ha resultado perjudicial, tales como la dependencia a este tipo de herramientas, el riesgo de plagio, la pérdida del pensamiento crítico y el riesgo a la generación de información falsa.

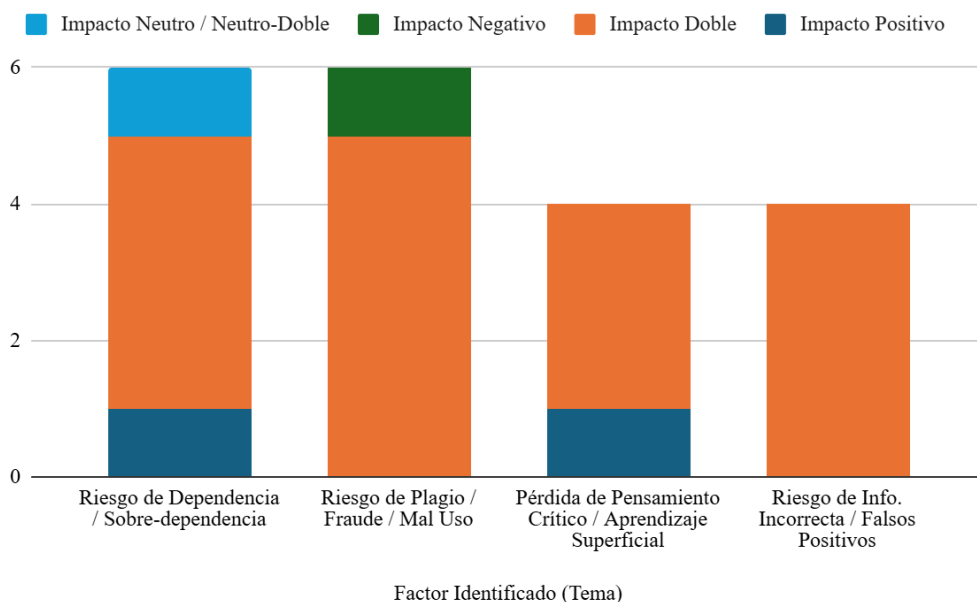


Figura 8. Impacto de los factores negativos en el aprendizaje

Los resultados obtenidos fueron:

- Entre los impactos negativos que conlleva la dependencia a este tipo de herramientas, se puede notar que aquellos estudiantes que recurrieron al uso de un asistente de inteligencia artificial, con frecuencia fueron menos eficientes y efectivos en el desarrollo de sus actividades (Mezzaro et al., 2024).

- La dependencia a este tipo de herramientas, en si no es mala, si su uso es guiado puede permitir aumentar la autoeficacia del estudiante, haciendo que la herramienta sea percibida como un apoyo para completar tareas complejas (Azoulay et al., 2025; Wang et al., 2025).
- La preocupación de que los estudiantes utilicen LLMs para generar soluciones ya hechas (Azoulay et al., 2025), obliga a los educadores a modificar sus prácticas de evaluación, lo cual se toma como un impacto positivo.
- El riesgo a posibles fraudes ha llevado a implementar métodos y herramientas de detección de trampas (Dong et al., 2025), y exámenes orales (Wang et al., 2023; Azoulay et al., 2025), lo que cambia el enfoque de evaluación del resultado final hacia la comprensión genuina del estudiante (Azoulay et al., 2025).
- En caso del pensamiento crítico, la IA se puede utilizar de apoyo para que los estudiantes revisen, critiquen y mejoren las soluciones generadas (Ata et al., 2025). Dicha práctica de evaluación crítica se convierte en una habilidad clave que desarrolla el pensamiento crítico (Azoulay et al., 2025).
- La capacidad que tienen los LLMs de generar información falsa y respuestas que pueden causar que los estudiantes desarrollen modelos mentales defectuosos (Wang et al., 2023).
- La generación de contenido errado por una IA puede convertirse en una oportunidad para mejorar la experiencia de aprendizaje (Wang et al., 2023).
- Todos los factores negativos generaron un impacto doble en algunos artículos

Basado en los resultados obtenidos anteriormente se pueden clasificar cada uno de los factores de impacto según su nivel de importancia, determinado por la correlación entre el factor y el tipo de impacto que generó en los distintos estudios.

Tabla 5. Nivel de importancia de los factores de impacto analizados

Factor de impacto	Nivel de importancia	Tipo
Asistencia en tareas, depuración y eficiencia	Muy alto	Positivo
Aumento de motivación	Alto	Positivo
Desarrollo de habilidades de orden superior	Dual/Muy alto	Positivo
Reducción de carga cognitiva	Alto	Positivo
Riesgo de dependencia o sobre dependencia	Crítico/Muy alto	Negativo
Riesgo de plagio o fraude	Crítico/Muy alto	Negativo
Pérdida de pensamiento crítico	Crítico/Muy alto	Negativo
Riesgo de información incorrecta o falsos positivos	Alto	Negativo

Conclusiones

El uso de herramientas de inteligencia artificial en la enseñanza de programación ha demostrado impactos positivos; herramientas como Chat GPT y Google Copilot permiten aumentar la eficiencia en la resolución de problemas, la motivación y confianza, debido a que ofrecen retroalimentación inmediata y asistencia personalizada a sus usuarios. Adicionalmente, contribuyen con el desarrollo de habilidades cognitivas, tales como el pensamiento, la resolución de problemas y la reducción de la carga cognitiva, siempre y cuando su uso sea asistido por un docente.

Los estudios revisados evidencian mayoritariamente un impacto positivo en la educación de la programación, por tanto, son herramientas que potencian el aprendizaje y las habilidades de los estudiantes. Sin embargo, un uso no regulado o sin acompañamiento pedagógico puede generar algunos impactos negativos, los cuales se encaminan a generar dependencia excesiva, realizar plagio y la pérdida del pensamiento crítico. Estos riesgos se originan cuando este tipo de herramientas dejan de usarse como un apoyo al aprendizaje. Por dicha razón, el uso de la IA en la educación genera impactos duales (positivos y negativos).

Referencias

- Abdulla, S., Ismail, S., Fawzy, Y., Elhag, A. (2024). Using ChatGPT in Teaching Computer Programming and Studying its Impact on Students Performance. *The Electronic Journal of e-Learning*, 22(6), 66-81. <https://doi.org/10.34190/ejel.22.6.3380>
- Atta, A., Esmat, M., Amasha, N., Elayat, E., ElSaid, W. K. (2025). A Smart ChatGPT Mobile Application for Improving C# Programming Skills for Students in Educational Institutions. *Qubahan Academic Journal*, 5(2), 49-62. <https://doi.org/10.48161/qaj.v5n2a1772>
- Azoulay, R., Hirst, T., Reches, S. (2025). Large Language Models in Computer Science Classrooms: Ethical Challenges and Strategic Solutions. *Applied Sciences*, 15(4), e1793. <https://doi.org/10.3390/app15041793>
- Baláz, N., Porubán, J., Horváth, M., Kormaník, T. (2024). Using ChatGPT During Implementation of Programs in Education. En *14th International Conference on Program Comprehension in Education (ICPEC 2024)* (Vol. 119, Art. 18). <https://doi.org/10.4230/oasics.icpec.2024.18>
- Bosch Global (s. f.). *Inteligencia artificial: los impactos de la tecnología que ya está cambiando el mundo*. <https://www.bosch.com.co/noticias-e-historias/aiot/inteligencia-artificial-los-impactos-de-la-tecnologia-que-ya-esta-cambiando-el-mundo/>
- Chang, C., Choi, W., Choi, I., Lei, H. (2024). A Systematic Literature Review of the Practical Applications of Artificial Intelligence-Generated Content (AIGC) Using OpenAI ChatGPT, Copilot, and Codex in Programming Education. En *Proceedings of the 4th International Conference on Artificial Intelligence and Virtual Reality* (pp. 13-19). <https://doi.org/10.1145/3719487.3719519>

Impacto de la IA en la enseñanza de la programación: Una revisión sistemática

- Dong, X., Wang, Z., Han, S. (2025). Mitigating Learning Burnout Caused by Generative Artificial Intelligence Misuse in Higher Education: A Case Study in Programming Language Teaching. *Informatics*, 12(2), e51. <https://doi.org/10.3390/informatics12020051>
- Ferino, S., Hoda, R., Grundy, J., Treude, C. (2025). Novice Developers' Perspectives on Adopting LLMs for Software Development: A Systematic Literature Review. *arXiv*. <https://doi.org/10.48550/arxiv.2503.07556>
- Gaitantzi, A., Kazanidis, I. (2025). The Role of Artificial Intelligence in Computer Science Education: A Systematic Review with a Focus on Database Instruction. *Applied Sciences*, 15(7), e3960. <https://doi.org/10.3390/app15073960>
- Kharrufa, A., Alghamdi, S., Aziz, A., Bull, C. (2024). LLMs Integration in Software Engineering Team Projects: Roles, Impact, and a Pedagogical Design Space for AI Tools in Computing Education. *arXiv*. <https://doi.org/10.48550/arxiv.2410.23069>
- Kosar, T., Ostojčić, D., Liu, Y. D., Mernik, M. (2024). Computer Science Education in ChatGPT Era: Experiences from an Experiment in a Programming Course for Novice Programmers. *Mathematics*, 12(5), e629. <https://doi.org/10.3390/math12050629>
- Li, J., Meland, P. H., Notland, J. S., Storhaug, A., Tysse, J. H. (2023). Evaluating the Impact of ChatGPT on Exercises of a Software Security Course. *ESEM, 2023*, 1-6. <https://doi.org/10.1109/esem56168.2023.10304857>
- Llerena-Izquierdo, J., Mendez-Reyes, J., Ayala-Carabajo, R., Andrade-Martinez, C. (2024). Innovations in Introductory Programming Education: The Role of AI with Google Colab and Gemini. *Education Sciences*, 14(12), e1330. <https://doi.org/10.3390/educsci14121330>
- Lyu, W., Wang, Y., Tingting, Chung, Sun, Y., Zhang, Y. (2024). Evaluating the Effectiveness of LLMs in Introductory Computer Science Education: A Semester-Long Field Study. *En Proceedings of the Eleventh ACM Conference on Learning @ Scale* (pp. 63-74). <https://doi.org/10.1145/3657604.3662036>
- Ma, B., Chen, L., Konomi, S. (2024). Enhancing Programming Education with ChatGPT: A Case Study on Student Perceptions and Interactions in a Python Course. *En Communications in Computer and Information Science* (pp. 113-126). https://doi.org/10.1007/978-3-031-64315-6_9
- Mellado, R., Cubillos, C. (2025). Can Generative Artificial Intelligence Outperform Self-Instructional Learning in Computer Programming?: Impact on Motivation and Knowledge Acquisition. *Applied Sciences*, 15(11), e5867. <https://doi.org/10.3390/app15115867>
- Mezzaro, S., Gambi, A., Fraser, G. (2024). An Empirical Study on How Large Language Models Impact Software Testing Learning. *En 28th International Conference on Evaluation and Assessment in Software Engineering, EASE 2024* (pp. 555-564). <https://doi.org/10.1145/3661167.3661273>
- Nathaniel, J., Oyelere, S., Suhonen, J., Tedre, M. (2025). Investigating the impact of generative AI integration on the sustenance of higher-order thinking skills and understanding of programming logic. *Computers and Education Artificial Intelligence*, 9, e100460. <https://doi.org/10.1016/j.caeai.2025.100460>
- Oyelere, S., Aruleba, K. (2025). A comparative study of student perceptions on generative AI in programming education across Sub-Saharan Africa. *Computers and Education Open*, 8, e100245. <https://doi.org/10.1016/j.caeo.2025.100245>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., *et al.* (2021). Declaración PRISMA 2020: una guía actualizada para la publicación de revisiones sistemáticas. *Revista Española de Cardiología*, 74(9), 790-799. <https://doi.org/10.1016/j.recesp.2021.06.016>
- PRISMA. (s. f.). *PRISMA Statement*. <https://www.prisma-statement.org/>

- PuntoEdu. (2016). *¿Por qué investigar en inglés hoy?*. <https://puntoedu.pucp.edu.pe/investigaciones/investigacion/por-que-investigar-en-ingles-hoy/>
- Sunday, K., Oyelere, S. S., Agbo, F. J., Aliyu, M. B., Balogun, O. S., Bouali, N. (2022). Usability Evaluation of Imikode Virtual Reality Game to Facilitate Learning of Object-Oriented Programming. *Technology Knowledge and Learning*, 28(4), 1871-1902. <https://doi.org/10.1007/s10758-022-09634-6>
- Velázquez, A. (2023). *¿Qué es la investigación empírica?*. <https://www.questionpro.com/blog/es/investigacion-empirica/>
- Vizcaino, A. (2020). *Inglés: El lenguaje de la ciencia* [Conjunto de datos]. <https://doi.org/10.3916/escuela-de-autores-139>
- Wang, H., Wang, C., Chen, Z., Liu, F., Bao, C., Xu, X. (2025). Impact of AI-agent-supported collaborative learning on the learning outcomes of University programming courses. *Education and Information Technologies*, 30(12), 17717-17749. <https://doi.org/10.1007/s10639-025-13487-8>
- Wang, T., Díaz, D. V., Brown, C., Chen, Y. (2023). Exploring the Role of AI Assistants in Computer Science Education: Methods, Implications, and Instructor Perspectives. En *IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)* (pp. 92-102). <https://doi.org/10.1109/vl-hcc57772.2023.00018>
- ATLAS. (2025). *¿Qué es una revisión sistemática de la literatura?*. <https://atlasti.com/es/guias/revisiones-bibliograficas/revision-sistemica>

Anexos

Anexo 1- Características de los artículos

Autor(es)	Año	Medio de publicación o Fuente	País/Región de realización	Tipo de estudio
Li et al.	2023	IEEE Xplore	Trondheim, Noruega	Empírico
Wang et al.	2023	IEEE Xplore	—	Métodos mixtos
Atta et al.	2025	Qubahan Academic Journal	Salerno, Italia	Métodos mixtos
Azoulay et al.	2025	Appl. Sci. 2025	Israel	Revisión sistemática
Wang et al.	2025	Education and Information Technologies (2025)	China	Cuasi-experimental
Baláž et al.	2024	5th International Computer Programming Education Conference (ICPEC 2024)	Eslovaquia	Empírico
Sunday et al.	2023	5th International Computer Programming Education Conference (ICPEC 2024)	Eslovaquia	Métodos mixtos
Ma et al.	2024	AIED 2024 Workshops, CCIS 2150	Japón	Empírico
Kosar et al.	2024	Mathematics 2024	Eslovenia	Cuasi-experimental/Controlado
Abdulla et al.	2024	Electronic Journal of e-Learning	Omán	Métodos mixtos
Mezzaro et al.	2024	28th International Conference on Evaluation and Assessment in Software Engineering (EASE 2024)	Alemania	Empírico
Lyu et al.	2024	Proceedings of the Eleventh ACM Conference on Learning @ Scale (L@S '24)	Estados Unidos (USA)	Métodos mixtos
Llerena et al.	2024	Educ. Sci. 2024	Ecuador	Cuantitativo cuasi-experimental
Dong et al.	2025	Informatics 2025	China	Cuantitativo
Nathaniel et al.	2025	Computers and Education: Artificial Intelligence	Finlandia	Métodos mixtos
Oyelere et al.	2025	Computers and Education Open 8 (2025)	Kenia, Nigeria y Sudáfrica	Cuantitativo
Mellado et al.	2025	Appl. Sci. 2025	Chile	Cuasi-experimental
Chang et al.	2025	2024 8th International Conference on Education and E-Learning (ICEEL)	EEUU Y China	Revisión sistemática