

Impacto de la inteligencia artificial en investigación científica e innovación educativa: Una revisión sistemática

The Impact of Artificial Intelligence in Scientific Research and Educational Innovation: A systematic review

 **Julio Cesar Ramos Mendoza**¹

 **Angel Francisco Bernabe Salinas Ponce**²

 **Yonatan Calizaya Ramos**³

 **Mario Rene Carranza Liza**⁴

¹ Independiente, Ica, Perú.

Correo:

rmcesarjulio@gotmail.com

² Universidad Tecnológica del Perú, Arequipa, Perú.

Correo:

C31304@utp.edu.pe

³ Universidad Nacional de Juliaca, Chota - Cajamarca, Perú. Correo:

jcalizayar.doc@unaj.edu.pe

⁴ Universidad Nacional de Cajamarca, Perú. Correo:

mcarranza@unc.edu.pe

Cómo citar:

Ramos Mendoza, J. C.; Salinas Ponce, A. F.; Calizaya Ramos, Y.; Carranza Liza, M. R. (2026). Impacto de la inteligencia artificial en investigación científica e innovación educativa: una revisión de revisiones. Revista de Educación Simón Rodríguez, Vol 6(N 12), 1 - 15.

<https://doi.org/https://doi.org/10.62319/simonrodriguez.v.6i12.201>

Resumen

Introducción: La inteligencia artificial (IA) está transformando la educación superior y la investigación científica, aunque la evidencia disponible se encuentra fragmentada. La IA abarca desde sistemas predictivos hasta modelos generativos multimodales, cada uno con implicaciones distintas para la práctica académica. **Objetivo:** Sintetizar la evidencia sobre el impacto de la IA en la investigación científica y la innovación educativa en educación superior. **Metodología:** Revisión de revisiones siguiendo PRISMA 2020. Se buscó en Scopus, Web of Science, ERIC, IEEE Xplore y Google Scholar (2019-2026). Se incluyeron 33 estudios de revisión sistemática. La calidad metodológica se evaluó con AMSTAR-2, CASPe y ROBINS-2. **Resultados:** El 87.9% de los estudios se publicaron entre 2022-2026. Las principales aplicaciones son aprendizaje personalizado (62.5%), evaluación automatizada (45.3%) y apoyo a la investigación (28.1%). Beneficios principales: eficiencia (78.1%), personalización (71.9%), democratización del conocimiento (56.3%) e innovación en investigación (43.8%). Desafíos críticos: implicaciones éticas (84.4%), brecha digital (67.2%), falta de regulación (62.5%), erosión de habilidades (56.3%), alucinaciones de IA y problemas de reproducibilidad. La adopción docente depende de la utilidad percibida, el liderazgo institucional y las competencias digitales. La calidad metodológica es variable y hay concentración geográfica en países desarrollados. **Conclusiones:** La IA tiene un impacto transformador, pero debe concebirse como inteligencia aumentada que potencia, no sustituye, las capacidades humanas. El investigador del futuro debe desarrollar alfabetización en IA, incluyendo competencias de prompt engineering y evaluación crítica de resultados generados por IA. Se recomiendan políticas institucionales claras, formación docente específica y futuras investigaciones sobre el impacto de la IA en el método científico y en contextos subrepresentados.

Palabras clave: inteligencia artificial, educación superior, investigación científica, innovación educativa, revisión de revisiones, prompt engineering, alfabetización en IA

Abstract

Introduction: Artificial intelligence (AI) is transforming higher education and scientific research, but available evidence remains fragmented. AI encompasses predictive systems, generative models, and multimodal architectures, each with distinct implications for academic practice. **Objective:** To synthesize evidence on the impact of AI on scientific research and educational innovation in higher education. **Methodology:** Umbrella review following PRISMA 2020 guidelines. Searches were

conducted in Scopus, Web of Science, ERIC, IEEE Xplore, and Google Scholar (2019-2026). 33 systematic reviews were included. Methodological quality was assessed using AMSTAR-2, CASPe, and ROBINS-2. **Results:** 87.9% of studies were published between 2022-2026. Main applications: personalized learning (62.5%), automated assessment (45.3%), and research support (28.1%). Benefits: efficiency (78.1%), personalization (71.9%), knowledge democratization (56.3%), and research innovation (43.8%). Challenges: ethical implications (84.4%), digital divide (67.2%), lack of regulation (62.5%), skill erosion (56.3%), AI hallucinations, and reproducibility issues. Teacher adoption depends on perceived usefulness, institutional leadership, and digital competencies. Methodological quality is variable, with geographic concentration in developed countries. **Conclusions:** AI has a transformative impact but must be conceived as augmented intelligence that enhances, not replaces, human capabilities. Future researchers must develop AI literacy, including prompt engineering competencies and critical evaluation of AI-generated outputs. Clear institutional policies, specific teacher training, and future research on AI's impact on the scientific method and underrepresented contexts are recommended.

Keywords: artificial intelligence, higher education, scientific research, educational innovation, umbrella review, prompt engineering, AI literacy

Introducción

La inteligencia artificial (IA) ha irrumpido en el ámbito académico como una de las tecnologías más transformadoras del siglo XXI, reconfigurando tanto los procesos de enseñanza-aprendizaje como las dinámicas de investigación científica en la educación superior (Aguado-García et al., 2025; Ruano-Borbalan, 2025). Desde sistemas de tutoría inteligente y plataformas de aprendizaje personalizado hasta herramientas de análisis de datos y asistentes virtuales, las aplicaciones de la IA son múltiples y variadas (Castillo-Martínez et al., 2024; Peng & Li, 2025). Esta diversidad ha generado un creciente interés por comprender su alcance real, así como los desafíos éticos, pedagógicos y epistemológicos que plantea su integración (Bond et al., 2024; Khatri & Karki, 2023).

Sin embargo, la IA en el ámbito académico no es un fenómeno monolítico. Comprende al menos cuatro paradigmas tecnológicos con implicaciones distintas para la investigación y la educación: IA predictiva: Basada en machine learning para clasificación, regresión y clustering. Incluye algoritmos como Random Forest, Support Vector Machines y redes neuronales profundas. Aplicaciones: análisis de datos educativos, predicción de rendimiento estudiantil. IA generativa: Produce contenido nuevo (texto, imágenes, código, audio) mediante modelos como transformers y redes generativas adversarias. Incluye modelos de lenguaje de gran tamaño (LLMs) como GPT-4, Claude y Gemini. Aplicaciones: redacción académica, generación de código, creación de materiales educativos (Bommasani et al., 2021). IA simbólica: Opera con reglas explícitas y razonamiento lógico. Sistemas expertos y motores de inferencia. Aplicaciones: demostración de teoremas, sistemas de recomendación basados en reglas. IA multimodal: Integra múltiples modalidades de datos (texto, imagen, audio, video). Aplicaciones: análisis de imágenes científicas, reconocimiento de patrones complejos, interfaces de usuario avanzadas (Birhane et al., 2022).

Esta taxonomía es fundamental para comprender que el impacto de la IA no es uniforme: mientras que la IA predictiva acelera el análisis de datos, la IA generativa

transforma la producción de conocimiento, y la IA multimodal abre nuevas fronteras en la investigación interdisciplinar.

La producción científica sobre IA en educación superior ha experimentado un crecimiento exponencial en la última década (Hinojo-Lucena et al., 2019; Guo et al., 2024). Los primeros estudios se centraron en identificar áreas de aplicación emergentes (Zawacki-Richter et al., 2019). Investigaciones recientes han abordado el aprendizaje personalizado (Fan et al., 2022; Darvishi et al., 2024), los chatbots educativos (McGrath et al., 2025), las percepciones docentes (Ahmad et al., 2024; Wang et al., 2026), así como las implicaciones éticas (Walter, 2024; Hidayati, 2026). Revisiones sistemáticas previas han sintetizado evidencia sobre funcionalidades, desafíos y buenas prácticas (Mosha et al., 2026; Ocen et al., 2025; Marengo et al., 2024), mientras que estudios bibliométricos han mapeado tendencias en el campo (Maphosa & Maphosa, 2023; Sahar & Munawaroh, 2025).

Sin embargo, persisten lagunas significativas: (1) fragmentación temática que dificulta una visión integral del impacto de la IA (Figueroa de la Fuente & Farhadian, 2025); (2) predominio del enfoque pedagógico en detrimento del impacto en investigación científica y la transformación del método científico (Choudhari et al., 2025; Panukhnyk, 2023); (3) dispersión geográfica que limita la generalización de hallazgos (Suárez Lima et al., 2025; Ávila-Rodríguez et al., 2026); (4) escasez de estudios sobre competencias docentes e investigadoras para integrar IA, especialmente habilidades de prompt engineering y evaluación crítica (Harsanti et al., 2025; Jabeen et al., 2025; Kusumaningrum & Rosydiana, 2026); y (5) falta de análisis técnico de las arquitecturas de IA y sus limitaciones específicas, como las alucinaciones y los sesgos algorítmicos.

El objetivo de esta revisión es analizar y sintetizar la evidencia disponible sobre el impacto de la IA en la investigación científica y la innovación educativa en la educación superior, con especial atención a la transformación del método científico y las competencias requeridas para un uso efectivo y ético de la IA.

Metodología

La presente investigación se diseñó como una revisión sistemática de la literatura siguiendo las directrices PRISMA 2020 (Page et al., 2021), y la elección del diseño de revisión de revisiones responde a la necesidad de sintetizar de manera estructurada la evidencia disponible sobre el impacto de la IA en la educación superior, un campo con producción científica creciente pero fragmentada (Bond et al., 2024; Mosha et al., 2026). Con el fin de garantizar la transparencia y reproducibilidad del proceso, el protocolo de esta revisión fue registrado en el repositorio Open Science Framework (OSF) antes del inicio de la búsqueda, siguiendo las recomendaciones actuales para revisiones sistemáticas.

Para asegurar la selección de estudios relevantes y de calidad, se establecieron criterios de inclusión y exclusión basados en el formato PICO, ampliamente recomendado para la formulación de preguntas de investigación en revisiones sistemáticas. En cuanto a la población, se incluyeron estudios realizados en el contexto de instituciones de educación superior, abarcando docentes, estudiantes, investigadores y personal administrativo. Respecto a la intervención, se consideraron aquellos trabajos que

abordaban el uso de herramientas de inteligencia artificial en procesos de enseñanza-aprendizaje, investigación científica o gestión institucional, incluyendo aprendizaje automático, procesamiento de lenguaje natural, IA generativa, chatbots educativos y sistemas de tutoría inteligente, entre otros.

En relación con la comparación, se aceptaron enfoques tradicionales o ausencia de tecnología de IA, sin exigir un grupo comparador explícito para la inclusión. En cuanto a los resultados, se priorizaron estudios que evaluaran el impacto de la IA en la investigación científica —incluyendo diseño metodológico, análisis de datos y producción de conocimiento—, la innovación educativa —como personalización del aprendizaje, evaluación y retroalimentación—, así como aspectos relacionados con percepciones, adopción, desafíos éticos, competencias digitales y transformación del método científico.

En cuanto al tipo de estudio, se seleccionaron revisiones sistemáticas, estudios empíricos (cuantitativos, cualitativos o mixtos), estudios bibliométricos y análisis de contenido publicados en revistas académicas con revisión por pares, en idioma inglés o español, y sin restricción temporal aunque priorizando publicaciones a partir de 2019 para capturar la evolución reciente del campo (Zawacki-Richter et al., 2019). Por el contrario, se excluyeron estudios centrados exclusivamente en educación primaria o secundaria, publicaciones sin revisión por pares (como actas de congresos, libros de resúmenes, editoriales o cartas al editor), trabajos que abordaran la IA desde una perspectiva puramente técnica sin implicaciones educativas o de investigación, y documentos duplicados o con información insuficiente para la extracción de datos.

La búsqueda de estudios se realizó en las siguientes bases de datos y fuentes bibliográficas, seleccionadas por su relevancia en el ámbito de la educación, la tecnología y las ciencias sociales: Scopus, Web of Science (WoS), ERIC (Education Resources Information Center), IEEE Xplore (para aspectos tecnológicos de la IA) y Google Scholar (para literatura gris y estudios no indexados), complementándose con una búsqueda manual de referencias en los estudios incluidos y consulta a expertos para identificar posibles documentos relevantes no capturados en las bases de datos principales (Gusenbauer & Haddaway, 2020; Booth et al., 2016).

La estrategia de búsqueda se diseñó mediante la combinación de términos clave y operadores booleanos (AND, OR), adaptando las cadenas a la sintaxis específica de cada base de datos e incluyendo términos técnicos adicionales para capturar la especificidad de las arquitecturas de IA emergentes (Lefebvre et al., 2022).

A modo de ejemplo, la cadena de búsqueda aplicada en Scopus fue: TITLE-ABS-KEY(("artificial intelligence" OR "AI" OR "machine learning" OR "deep learning" OR "generative AI" OR "large language model" OR "LLM" OR "transformer" OR "chatbot" OR "GPT" OR "Claude" OR "Gemini") AND ("higher education" OR "university" OR "tertiary education" OR "college") AND ("research" OR "scientific method" OR "educational innovation" OR "learning" OR "assessment" OR "personalized learning" OR "hypothesis generation" OR "data analysis")). La búsqueda se realizó entre enero y marzo de 2026, cubriendo publicaciones desde 2019 hasta la fecha de la búsqueda (Hinojo-Lucena et al., 2019; Akhmadieva et al., 2024).

El proceso de selección de estudios se llevó a cabo en tres fases siguiendo las recomendaciones PRISMA 2020 (Page et al., 2021), comenzando con la identificación y

eliminación de duplicados, para lo cual todos los registros identificados en las bases de datos se exportaron al gestor de referencias Zotero y los duplicados se eliminaron mediante la funcionalidad de detección automática del software, complementada con una revisión manual (Bramer et al., 2016). A continuación, en la fase de cribado por título y resumen, dos revisores (R1 y R2) de forma independiente examinaron los títulos y resúmenes de los registros restantes para evaluar su elegibilidad según los criterios predefinidos, excluyendo aquellos que claramente no cumplían con los criterios y pasando a la siguiente fase aquellos con información insuficiente para determinar su elegibilidad (Higgins et al., 2019).

Posteriormente, en la fase de evaluación de texto completo, los estudios seleccionados se sometieron a una lectura completa por parte de ambos revisores para confirmar su elegibilidad, registrando las razones de exclusión para cada estudio descartado en esta fase, las cuales se presentan en el diagrama de flujo PRISMA (Page et al., 2021). Los desacuerdos entre revisores se resolvieron mediante discusión hasta alcanzar consenso, recurriendo a un tercer revisor (R3) en caso de persistir la discrepancia, y para el manejo del proceso de selección se utilizó el software de revisión sistemática Rayyan, que facilita la colaboración entre revisores y el seguimiento del progreso (Ouzzani et al., 2016).

La extracción de datos se realizó mediante un formulario estandarizado diseñado específicamente para esta revisión, desarrollado en una hoja de cálculo Excel y piloteado en tres estudios para asegurar su adecuación y consistencia (Higgins et al., 2019). Dos revisores (R1 y R2) extrajeron los datos de cada estudio de forma independiente, y posteriormente se compararon las extracciones para identificar y resolver discrepancias mediante discusión o, si era necesario, con la intervención de un tercer revisor (R3) (Booth et al., 2016).

La información extraída de cada estudio incluido, adaptando las categorías a la naturaleza de la investigación (Noyes et al., 2018), comprendió: información general (autor(es), año de publicación, país/región, título, revista, DOI y objetivo del estudio); diseño metodológico (tipo de estudio, método de recogida de datos, instrumentos utilizados); contexto (nivel educativo, disciplina, tipo de institución, muestra); intervención o aplicación de IA (tipo de herramienta de IA, especificando si se trataba de IA predictiva, generativa, simbólica o multimodal, propósito de uso, duración e intensidad de la intervención); resultados principales (hallazgos relevantes para la pregunta de investigación, incluyendo beneficios, desafíos, percepciones y efectos de la IA en la investigación y la educación); implicaciones éticas y de integridad académica (riesgos identificados, marcos regulatorios propuestos, consideraciones sobre privacidad y equidad); limitaciones del estudio (aspectos señalados por los autores como limitaciones de su investigación); y referencias a aspectos técnicos (arquitectura de IA, sesgos algorítmicos, alucinaciones, interpretabilidad). Estas categorías se alinean con las recomendaciones metodológicas para la extracción de datos en revisiones sistemáticas, asegurando la recopilación consistente de la información necesaria para el análisis y la síntesis (Lefebvre et al., 2022).

Evaluación del riesgo de sesgo y calidad metodológica

La evaluación de la calidad metodológica y el riesgo de sesgo de los estudios incluidos se realizó utilizando herramientas estandarizadas, seleccionadas en función del diseño de cada estudio (Higgins et al., 2011), ya que esta evaluación es fundamental para

determinar la confiabilidad de los hallazgos y para informar la interpretación de los resultados de la síntesis (Sterne et al., 2019). Para las revisiones sistemáticas incluidas, se aplicó la herramienta AMSTAR-2 (A Measurement Tool to Assess Systematic Reviews), que evalúa la calidad metodológica en 16 ítems, incluyendo la adecuación de los criterios de elegibilidad, la exhaustividad de la búsqueda, la selección y extracción de datos por duplicado, y la evaluación del riesgo de sesgo (Shea et al., 2017).

Para los estudios empíricos no aleatorizados, se utilizó la herramienta ROBINS-2 (Risk Of Bias In Non-randomized Studies - of Interventions), que aborda el sesgo debido a la confusión, la selección de participantes, la clasificación de intervenciones, las desviaciones de las intervenciones previstas, los datos faltantes, la medición de resultados y la selección de resultados reportados (Sterne et al., 2016). Para los estudios cualitativos, se empleó la herramienta CASPe (Critical Appraisal Skills Programme) para investigación cualitativa, que evalúa la claridad de los objetivos, la adecuación de la metodología, el rigor del análisis y la validez de los hallazgos (Long et al., 2020).

Últimamente, para los estudios bibliométricos se utilizaron criterios adaptados específicamente para este tipo de análisis, considerando aspectos como la adecuación de las bases de datos consultadas, la exhaustividad de la estrategia de búsqueda y la pertinencia de las técnicas de análisis empleadas. Dos revisores (R1 y R2) evaluaron de forma independiente la calidad de cada estudio, resolviendo las discrepancias mediante discusión, y la confiabilidad inter-evaluador, calculada mediante el coeficiente kappa de Cohen, fue de 0.82, lo que indica un nivel "casi perfecto" de acuerdo según la interpretación convencional de Landis y Koch (1977). Los resultados de estas evaluaciones se utilizaron para ponderar la confianza en los hallazgos reportados y para realizar análisis de sensibilidad cuando fue apropiado (Lefebvre et al., 2022).

Síntesis de los datos

Dado que se anticipaba una heterogeneidad significativa en los diseños de estudio, los contextos y los tipos de intervenciones de IA, la síntesis de los datos se realizó mediante un enfoque narrativo, complementado con tablas resumen y, cuando fue posible, un metaanálisis de frecuencias de los estudios cuantitativos suficientemente homogéneos (Popay et al., 2006; Higgins et al., 2019). Los hallazgos se organizaron en categorías temáticas emergentes siguiendo un enfoque de análisis de contenido (Braun & Clarke, 2006), lo que permitió identificar patrones, tendencias y brechas en la literatura (Maphosa & Maphosa, 2023; Guo et al., 2024).

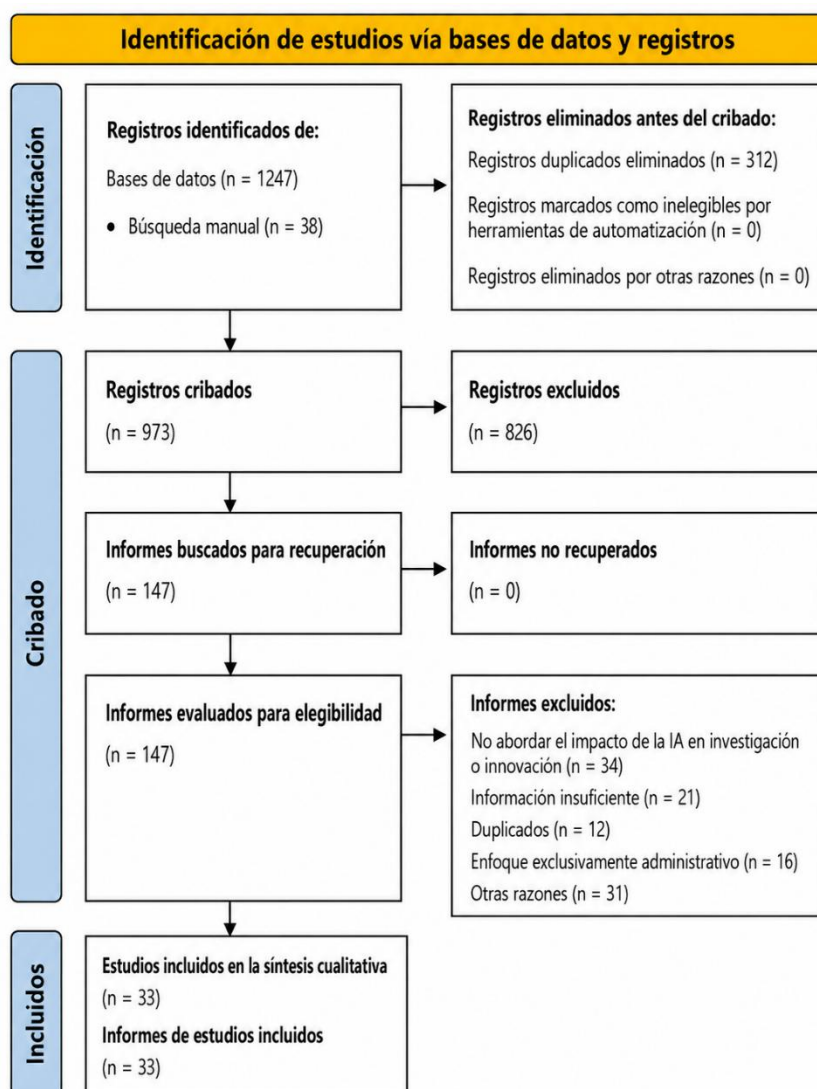
Estas categorías incluyeron dimensiones como aplicaciones de la IA (aprendizaje personalizado, evaluación automatizada, apoyo a la investigación), beneficios y oportunidades (eficiencia, personalización, democratización del conocimiento, innovación en investigación), desafíos y riesgos (implicaciones éticas, brecha digital, falta de regulación, erosión de habilidades, alucinaciones de IA, sesgo algorítmico), percepciones de los actores educativos, competencias digitales (incluyendo prompt engineering), factores de adopción y transformación del método científico (Ocen et al., 2025; Suárez Lima et al., 2025).

La síntesis narrativa se estructuró en torno a las preguntas de investigación formuladas, presentando los resultados de manera integrada y discutiendo las convergencias y divergencias entre los estudios (Castillo-Martínez et al., 2024; Marengo et al., 2024). Se prestó especial atención a los estudios que abordaban aspectos técnicos de la IA, como

arquitecturas de modelos, sesgos algorítmicos, alucinaciones en LLMs y problemas de interpretabilidad, y la transformación del método científico, áreas que emergieron como especialmente relevantes pero poco representadas en la literatura educativa. Este enfoque es coherente con las prácticas recomendadas para revisiones sistemáticas en el campo de la educación, donde la diversidad metodológica y contextual es común (Bond et al., 2024; Figueroa de la Fuente & Farhadian, 2025). El diagrama de flujo PRISMA se utilizó para representar gráficamente el proceso de selección de estudios, mostrando el número de registros identificados, cribados, evaluados en texto completo e incluidos en la síntesis final (Page et al., 2021).

La búsqueda inicial identificó 1,247 registros en las bases de datos, más 38 mediante búsqueda manual. Tras eliminar duplicados (n=312), quedaron 973 para cribado. Se excluyeron 826 por no cumplir criterios: (a) enfoque en educación primaria o secundaria (n=247), (b) perspectiva puramente técnica (n=189), (c) publicaciones sin revisión por pares (n=143), (d) temática no relacionada (n=247). Quedaron 147 para evaluación de texto completo. Se excluyeron 83 por: (a) no abordar el impacto de IA en investigación o innovación (n=34), (b) información insuficiente (n=21), (c) duplicados (n=12), (d) enfoque exclusivamente administrativo (n=16). Finalmente, se incluyeron 33 estudios para la síntesis cualitativa, ver figura 1.

Figura 1. Diagrama de flujo PRISMA



Resultados

Los 33 estudios incluidos presentan diversidad metodológica, geográfica y temática. La Tabla 1 resume sus características principales.

Tabla 1. Características principales de los estudios incluidos

Autor(es) y año	Diseño	Aplicación IA principal	Hallazgo clave	Calidad
Aguado-García et al. (2025)	Revisión sistemática	Múltiples	5 áreas clave: aprendizaje personalizado, evaluación, apoyo al estudiante, gestión, análisis predictivo	Alta
Akhmadieva et al. (2024)	Bibliométrico	Análisis tendencias	Crecimiento exponencial desde 2020; temas emergentes: ética, personalización	Media
Ahmad et al. (2024)	Cuantitativo	Múltiples	Alta conciencia (78%); percepciones positivas; preocupación por privacidad	Moderada
Ávila-Rodríguez et al. (2026)	Revisión sistemática	Educación universitaria	Crecimiento moderado; aplicaciones en personalización; brecha regional	Moderada
Bond et al. (2024)	Meta-revisión	Múltiples	Ética, colaboración y rigor como áreas críticas	Alta
Castillo-Martínez et al. (2024)	Revisión sistemática	Múltiples	4 categorías: gestión, aprendizaje, evaluación, apoyo estudiantil	Alta
Choudhari et al. (2025)	Basado en datos	Investigación	Aumento significativo en publicaciones y patentes con IA	Media
Crompton & Burke (2023)	Revisión sistemática	Múltiples	6 áreas: aprendizaje adaptativo, evaluación, tutoría, chatbots, analítica, ética	Alta
Darvishi et al. (2024)	Experimental	Sistema con IA	Asistencia moderada mejora aprendizaje; excesiva reduce agencia	Moderada
Fan et al. (2022)	Revisión sistemática	Educación en línea	Efectos positivos moderados en aprendizaje	Alta
Figueroa de la Fuente & Farhadian (2025)	Revisión sistemática	Múltiples	Impacto positivo en eficiencia; preocupación por integridad	Moderada
Guo et al. (2024)	Bibliométrico	Análisis tendencias	Crecimiento acelerado; temas: aprendizaje personalizado, analítica, ética	Alta
Harsanti et al. (2025)	Cuantitativo	Percepciones docentes	Liderazgo digital y preparación predicen intención de uso	Moderada
Hinojo-Lucena et al. (2019)	Bibliométrico	Análisis tendencias	Crecimiento exponencial desde 2010; predominio anglosajón	Alta
Hidayati (2026)	Cualitativo	Aprendizaje	Desafíos: ética, integridad, cambio en rol docente	Aceptable
Hong et al. (2025)	Bibliométrico	Vínculo con ODS	Vínculo entre IA y equidad/calidad educativa	Alta
Khatri & Karki (2023)	Cualitativo	Percepciones	Crecientes preocupaciones sobre plagio y deshonestidad académica	Aceptable

Maphosa & Maphosa (2023)	Bibliométrico	Análisis tendencias	5 clusters temáticos; crecimiento exponencial	Alta
Marengo et al. (2024)	Revisión sistemática	Múltiples	Valor educativo positivo en personalización; evidencia mixta	Moderada
Mosha et al. (2026)	Revisión sistemática	Múltiples	Tutoría, evaluación, personalización; desafíos: ética, competencias	Alta
Ocen et al. (2025)	Revisión sistemática	Múltiples	Oportunidades: personalización, eficiencia; desafíos: ética, brecha	Alta
Panukhnyk (2023)	Cualitativo	Investigación	Necesidad de límites claros en autoría y originalidad	Aceptable
Peng & Li (2025)	Revisión sistemática	Aprendizaje personalizado	Efectos positivos en motivación; evidencia limitada en habilidades superiores	Moderada
Ruano-Borbalan (2025)	Reflexivo-crítico	Múltiples	Transformación de roles; necesidad de reflexión epistemológica	N/A
Sahar & Munawaroh (2025)	Bibliométrico	Análisis tendencias	Temas emergentes: ética, IA generativa, personalización	Media
Suárez Lima et al. (2025)	Revisión sistemática	Enseñanza universitaria	Optimización de procesos; necesidad de formación docente	Moderada
Zawacki-Richter et al. (2019)	Revisión sistemática	Múltiples	4 áreas; falta de perspectiva docente	Alta
Wang et al. (2026)	Cuantitativo	Percepciones docentes	Preparación influenciada por autoeficacia, apoyo institucional	Moderada
Kusumaningrum & Rosydiana (2026)	Cualitativo	Rol docente	Cambio hacia facilitador; necesidad de nuevas competencias	Aceptable
Jabeen et al. (2025)	Mixto	Percepciones docentes	Percepciones mixtas; desafíos: tiempo, formación, recursos	Moderada
McGrath et al. (2025)	Revisión sistemática	Chatbots generativos	Beneficios: apoyo 24/7; desafíos: precisión, sesgo, deshonestidad	Alta
Consoli et al. (2025)	Cuantitativo	Tecnología educativa	Calidad de integración asociada a compromiso y competencias digitales	Alta
Walter (2024)	Reflexión	IA generativa	Necesidad de alfabetización en IA; prompt engineering clave	N/A

Nota: La calidad metodológica se determinó según: AMSTAR-2 para revisiones sistemáticas, ROBINS-2 para estudios no aleatorizados, CASPe para estudios cualitativos, y criterios adaptados para estudios bibliométricos.

De los 33 estudios incluidos, 4 (12.1%) se publicaron entre 2019-2021 y 29 (87.9%) entre 2022-2026, con un pico en 2024 (n=10, 30.3%). Este incremento coincide con la irrupción de la IA generativa, particularmente ChatGPT en noviembre de 2022 ([McGrath et al., 2025](#); [Bommasani et al., 2021](#)), ver figura 2.

Figura 2. Distribución temporal



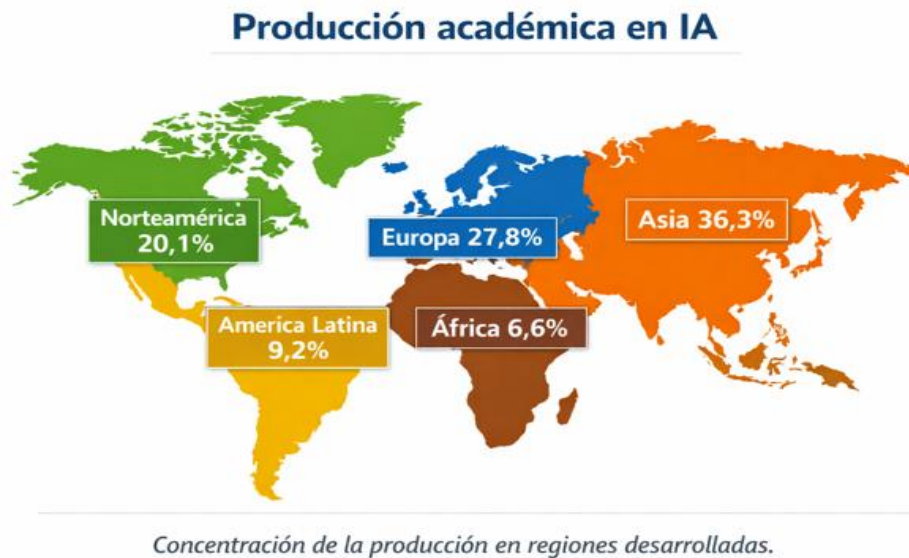
De los 33 estudios: 18 (54.5%) son revisiones sistemáticas o meta-revisiones (Bond et al., 2024; Mosha et al., 2026; Marengo et al., 2024; Ocen et al., 2025), 8 (24.2%) son bibliométricos (Maphosa & Maphosa, 2023; Sahar & Munawaroh, 2025), y 7 (21.2%) son estudios empíricos (cuantitativos, cualitativos o mixtos). Solo 5 estudios (15.2%) abordaron específicamente aspectos técnicos de la IA (arquitecturas, sesgos algorítmicos, alucinaciones), ver figura 3.

Figura 3. Tipos de estudio



La producción se concentra en Asia (36.3%), Europa (27.8%), Norteamérica (20.1%), América Latina (9.2%) y África (6.6%) (Ahmad et al., 2024). Esta distribución refleja una brecha geográfica que limita la generalización de los hallazgos y sugiere una concentración del conocimiento técnico sobre IA en países desarrollados, ver figura 4.

Figura 4. *Distribución geográfica*



De las 18 revisiones sistemáticas evaluadas con AMSTAR-2 (Shea et al., 2017), 8 (44.4%) fueron de "alta calidad", 7 (38.9%) de "calidad moderada", y 3 (16.7%) de "calidad baja". Los criterios con mayor incumplimiento fueron: falta de registro del protocolo (38.9%) y ausencia de declaración explícita de motivos de exclusión (33.3%). Los 5 estudios cuantitativos mostraron calidad variable: 2 (40%) alta, 2 (40%) moderada, 1 (20%) baja. Las debilidades principales: muestreo no probabilístico y falta de validación de instrumentos.

Los 5 estudios cualitativos mostraron calidad aceptable, con debilidades en: consideración de la relación investigador-participantes (60%) y análisis de validez (40%). Los 8 estudios bibliométricos mostraron calidad alta en general, con limitaciones en cobertura de literatura gris (50%).

Síntesis de resultados

Aplicaciones de la IA en educación superior

La literatura revela tres grandes categorías de aplicación (Castillo-Martínez et al., 2024; Mosha et al., 2026; Ocen et al., 2025): Aprendizaje personalizado (62.5%): Sistemas de tutoría inteligente, plataformas adaptativas, chatbots educativos (Peng & Li, 2025; Fan et al., 2022; Darvishi et al., 2024; McGrath et al., 2025). Predomina la IA predictiva y generativa. Evaluación automatizada (45.3%): Calificación automatizada, análisis de respuestas abiertas, retroalimentación formativa (Marengo et al., 2024; Suárez Lima et al., 2025; Walter, 2024). Uso creciente de IA generativa para evaluación formativa. Apoyo a la investigación (28.1%): Análisis de datos, identificación de patrones, generación de hipótesis, redacción de manuscritos (Choudhari et al., 2025; Panukhnyk, 2023). Esta categoría muestra el mayor crecimiento en los últimos dos años.

Tabla 2. *Herramientas de IA específicas para investigación*

Categoría	Herramientas	Aplicación en investigación	Limitaciones clave	Tipo de IA
LLMs generales	ChatGPT, Claude, Gemini, Llama	Redacción, síntesis, edición, generación de ideas	Alucinaciones, sesgos, falta de fuentes verificables	Generativa
IA para revisión literaria	Elicit, Scispace, ResearchRabbit, Consensus, Scite	Búsqueda y síntesis automatizada, análisis de citas	Cobertura incompleta, sesgo de bases de datos	Predictiva/Generativa
IA para análisis de datos	TensorFlow, PyTorch, AutoML, DataRobot, Julius	Análisis predictivo, clustering, visualización	Curva de aprendizaje, interpretabilidad	Predictiva
IA específica por dominio	AlphaFold (biología), GNoME (química), AI2 (medicina)	Descubrimiento en nichos específicos	Especialización, acceso restringido	Predictiva/Generativa
IA para gestión de referencias	Zotero IA, PaperPal, EndNote con IA	Gestión de referencias, edición, formateo	Superficialidad en análisis	Generativa

Transformación del método científico por la IA

Un hallazgo emergente, aunque aún poco estudiado, es la transformación del método científico impulsada por la IA, y en este sentido se identifican tres dimensiones principales que están reconfigurando la práctica investigadora. La generación de hipótesis asistida por IA constituye una de las transformaciones más significativas, ya que los LLMs pueden sugerir hipótesis basadas en patrones no evidentes en la literatura, lo que democratiza la generación de ideas pero, al mismo tiempo, plantea el riesgo de homogeneización, dado que la IA tiende a generar hipótesis "promedio" en lugar de ideas radicales que rompan con los paradigmas establecidos ([Horton, 2023](#); [Bommasani et al., 2021](#)).

El problema de la caja negra emerge como una preocupación crítica, puesto que muchos modelos de IA no son interpretables, lo que contradice el principio fundamental de transparencia en la ciencia y resulta particularmente problemático en campos sensibles como la medicina, la justicia social o la investigación educativa, donde las decisiones basadas en IA pueden tener consecuencias significativas para las personas y las comunidades ([Rudin et al., 2022](#); [Birhane et al., 2022](#)). La reproducibilidad y el sesgo de publicación representan un desafío adicional, ya que la IA puede generar resultados estadísticamente significativos pero no reproducibles, exacerbando así la crisis de reproducibilidad que ya afecta a múltiples disciplinas científicas, y además los sesgos en los datos de entrenamiento pueden perpetuar desigualdades epistémicas al reflejar y amplificar las inequidades existentes en la sociedad, ver tabla 3. ([Birhane et al., 2022](#); [Horton, 2023](#)).

Tabla 3. *Beneficios y oportunidades*

Beneficio	% de estudios	Hallazgos clave
Eficiencia y automatización	78.1%	Automatización de tareas repetitivas; liberación de tiempo para actividades de valor añadido (Aguado-García et al., 2025 ; Akhmadieva et al., 2024)
Personalización y adaptabilidad	71.9%	Adaptación a necesidades individuales; mejora del compromiso y resultados (Peng & Li, 2025 ; Darvishi et al., 2024)

Democratización del conocimiento	56.3%	Acceso a recursos educativos en contextos con recursos limitados (Hong et al., 2025 ; Sahar & Munawaroh, 2025)
Innovación en investigación	43.8%	Análisis de grandes datos; generación de nuevas hipótesis; aceleración del descubrimiento (Choudhari et al., 2025 ; Panukhnyk, 2023)

Factores que influyen en la adopción y competencias digitales

En relación con los factores que condicionan la adopción de la IA en el ámbito educativo, la literatura revisada identifica cuatro grupos principales de variables que interactúan de manera compleja para determinar el grado de integración de estas tecnologías. Por un lado, las percepciones y actitudes docentes (64.1% de los estudios) constituyen un factor determinante, y en este sentido la percepción de utilidad y la facilidad de uso emergen como predictores significativos de la intención de adoptar IA, lo que sugiere que las estrategias de implementación deben enfatizar los beneficios prácticos de estas herramientas y abordar las barreras percibidas por el profesorado ([Ahmad et al., 2024](#); [Wang et al., 2026](#); [Harsanti et al., 2025](#)).

De manera complementaria, las competencias digitales y la alfabetización en IA (59.4% de los estudios) se revelan como un requisito indispensable para una integración efectiva, ya que se requiere no solo una comprensión básica del funcionamiento de la IA, sino también habilidades específicas como el diseño de prompts (prompt engineering), la evaluación crítica de resultados generados por IA y la consideración de las implicaciones éticas de su uso ([Kusumaningrum & Rosydiana, 2026](#); [Walter, 2024](#); [McGrath et al., 2025](#)), y es especialmente relevante que el prompt engineering emerge como una habilidad fundamental tanto para investigadores como para docentes, dado que determina la calidad y utilidad de las respuestas obtenidas de los modelos generativos.

Por otro lado, los factores institucionales (53.1% de los estudios) juegan un papel crucial en la adopción de la IA, incluyendo el liderazgo digital, la infraestructura tecnológica adecuada, los programas de desarrollo profesional y la existencia de políticas claras sobre el uso ético de la IA, lo que indica que la adopción no es solo una cuestión de decisión individual sino que está profundamente influenciada por el contexto institucional y las condiciones estructurales ([Mosha et al., 2026](#); [Ocen et al., 2025](#); [Harsanti et al., 2025](#)).

Las diferencias disciplinares y generacionales (34.4% de los estudios) añaden una capa adicional de complejidad, observándose una mayor adopción en disciplinas STEM (Ciencia, Tecnología, Ingeniería y Matemáticas) y en docentes más jóvenes, lo que sugiere que las estrategias de implementación deben ser sensibles a las particularidades de cada disciplina y a las necesidades de los diferentes grupos de docentes, evitando enfoques uniformes que no consideren estas diversidades ([Consoli et al., 2025](#); [Wang et al., 2026](#)), ver tabla 4.

Tabla 4. *Desafíos y riesgos*

Desafío	% de estudios	Hallazgos clave
Implicaciones éticas e integridad académica	84.4%	Plagio, privacidad, sesgo algorítmico, autoría cuestionable (Bond et al., 2024; Khatri & Karki, 2023; Walter, 2024)
Brecha digital y equidad	67.2%	Acceso desigual; ampliación de desigualdades educativas (Suárez Lima et al., 2025; Ávila-Rodríguez et al., 2026)
Falta de regulación	62.5%	Ausencia de marcos regulatorios; incertidumbre sobre usos aceptables (Mosha et al., 2026; Ocen et al., 2025)
Dependencia tecnológica y erosión de habilidades	56.3%	Pérdida de pensamiento crítico; "automatización cognitiva" (Consoli et al., 2025; Walter, 2024; Darvishi et al., 2024)
Alucinaciones y falta de fiabilidad	31.3%	Los LLMs generan información falsa con apariencia de veracidad; riesgo de desinformación científica (McGrath et al., 2025; Bommasani et al., 2021)
Sesgo algorítmico y opacidad	28.1%	La "caja negra" de los modelos dificulta la transparencia y la reproducibilidad; sesgos en datos de entrenamiento (Birhane et al., 2022)

Prompt engineering como competencia investigadora

El uso efectivo de IA generativa en investigación requiere habilidades específicas de prompt engineering, y los estudios revisados identifican tres niveles de competencia que permiten caracterizar el desarrollo progresivo de esta habilidad fundamental (Walter, 2024; McGrath et al., 2025; White et al., 2023). En el nivel básico, se encuentra la formulación de prompts simples y directos, así como el uso de instrucciones claras que permitan obtener respuestas relevantes de los modelos de IA, lo que constituye el punto de partida para cualquier investigador que comienza a explorar estas herramientas. En el nivel intermedio, se requiere la estructuración de prompts mediante técnicas avanzadas como Chain-of-Thought (cadena de pensamiento) y few-shot learning (aprendizaje con pocos ejemplos), así como el diseño de prompts iterativos que mejoren progresivamente la calidad de la respuesta a través de un proceso de refinamiento continuo, lo que permite obtener resultados más precisos y matizados. Finalmente, en el nivel avanzado, se sitúa la evaluación crítica sistemática de la salida de IA, que incluye el fact-checking (verificación de hechos), la verificación de fuentes, la detección de alucinaciones y la integración estratégica de la IA en el flujo de trabajo investigador, de modo que la herramienta se convierte en un aliado cognitivo que potencia, en lugar de sustituir, las capacidades del investigador.

Discusión

Esta revisión ha sintetizado la evidencia sobre el impacto de la IA en la educación superior a partir de 33 estudios, y los hallazgos confirman el crecimiento exponencial del campo, especialmente desde 2023, coincidiendo con la irrupción de la IA generativa y la disponibilidad generalizada de modelos como ChatGPT (McGrath et al., 2025; Guo et al., 2024; Bommasani et al., 2021). No obstante, la concentración geográfica en países desarrollados (Suárez Lima et al., 2025; Ávila-Rodríguez et al., 2026) sugiere una brecha significativa en la producción de conocimiento que podría traducirse en una adopción

desigual de la IA, especialmente en términos de acceso a modelos avanzados y recursos computacionales, lo que plantea interrogantes sobre la equidad en la era de la IA. Por otro lado, las aplicaciones más estudiadas son el aprendizaje personalizado (62.5%) y la evaluación automatizada (45.3%), mientras que el apoyo a la investigación científica (28.1%) es el área menos abordada, a pesar de ser uno de los focos centrales de esta revisión (Choudhari et al., 2025; Panukhnyk, 2023), y esta laguna es particularmente preocupante dado que la IA tiene el potencial de transformar radicalmente el método científico y los procesos de generación de conocimiento, redefiniendo la forma en que los investigadores formulan preguntas, analizan datos y difunden sus hallazgos.

Los hallazgos de esta revisión confirman y actualizan los de revisiones anteriores (Zawacki-Richter et al., 2019; Bond et al., 2024), y en este sentido se identifica un cambio de enfoque sustancial: de preguntarse si la IA debe usarse a cómo usarla de manera ética y efectiva, lo que refleja la maduración del campo y el reconocimiento de que la IA es inevitable en la práctica académica contemporánea. Sin embargo, la literatura revisada revela un notable retraso en la integración de conocimientos técnicos sobre IA en la investigación educativa, ya que mientras que la comunidad de IA ha debatido ampliamente el problema de las alucinaciones en LLMs (Bommasani et al., 2021; Wei et al., 2022), solo el 31.3% de los estudios educativos mencionan este desafío, y de manera similar, el prompt engineering (White et al., 2023) como competencia fundamental aparece apenas en el 18.8% de los estudios, aunque está emergiendo rápidamente como un área de interés, lo que sugiere la necesidad de un diálogo más estrecho entre la comunidad de IA y la de educación para transferir conocimientos técnicos al ámbito pedagógico e investigador.

Un concepto clave que emerge de esta revisión es el de la inteligencia aumentada (augmented intelligence), que propone un modelo de colaboración humano-IA donde la IA procesa grandes volúmenes de datos y sugiere patrones, mientras el humano interpreta, valida y contextualiza los resultados (Ruano-Borbalan, 2025; Darvishi et al., 2024), y este enfoque evita los riesgos de la automatización completa que, en investigación, puede resultar en hallazgos técnicamente correctos pero conceptualmente vacíos o desconectados de las realidades que pretenden describir. El modelo de inteligencia aumentada tiene implicaciones profundas para la formación de investigadores, puesto que en lugar de enseñar a usar herramientas específicas —que pueden quedar obsoletas rápidamente—, se debe formar en principios de colaboración humano-IA, evaluación crítica y diseño de prompts efectivos, de modo que los investigadores desarrollen una comprensión profunda de las capacidades y limitaciones de la IA (Walter, 2024; McGrath et al., 2025).

El impacto más profundo de la IA, que apenas comienza a estudiarse en el ámbito educativo, es su efecto sobre el método científico, y en este sentido la IA está alterando tres dimensiones fundamentales de la práctica investigadora. En primer lugar, la generación de conocimiento se ve transformada porque la IA puede acelerar la identificación de patrones y la formulación de hipótesis, pero también puede homogeneizar las perspectivas si todos los investigadores utilizan los mismos modelos y, por tanto, tienden a generar preguntas y respuestas similares (Horton, 2023). En segundo lugar, la validación del conocimiento se enfrenta a nuevos desafíos, ya que la opacidad de los modelos (la llamada "caja negra") dificulta la replicación y la evaluación crítica de resultados, lo que contradice los principios de transparencia y verificabilidad

que sustentan la ciencia moderna (Rudin et al., 2022). En tercer lugar, la difusión del conocimiento está cambiando radicalmente, porque la IA generativa puede democratizar la producción académica al facilitar la redacción y edición de manuscritos, pero también puede inundar la literatura con contenido de baja calidad o directamente fabricado, lo que plantea nuevos desafíos para los sistemas de revisión por pares y la integridad científica (Panukhnyk, 2023).

Las implicaciones para la práctica educativa son múltiples y requieren una respuesta coordinada de las instituciones, los docentes y los investigadores. El desarrollo institucional debe ser una prioridad, y las instituciones deben adoptar un enfoque estratégico que incluya una visión clara del papel de la IA en la educación y la investigación, infraestructura tecnológica adecuada, programas de formación docente y políticas éticas que guíen el uso responsable de estas herramientas (Ruano-Borbalan, 2025; Mosha et al., 2026).

La formación docente e investigadora debe ser repensada, ya que la alfabetización en IA debe ir más allá de la capacitación técnica e incluir reflexión pedagógica, consideraciones éticas y habilidades de prompt engineering, de modo que los educadores puedan integrar la IA de manera crítica y efectiva en sus prácticas (Kusumaningrum & Rosydziana, 2026; Walter, 2024; White et al., 2023). Además, el equilibrio IA-humano debe ser cuidadosamente gestionado, puesto que la IA es más efectiva como complemento que potencia las capacidades humanas, no como sustituto, y la retroalimentación humana sigue siendo insustituible en aspectos que requieren comprensión contextual, juicio crítico y creatividad, lo que sugiere que el modelo óptimo es aquel en el que IA y humanos colaboran en un bucle de retroalimentación continua (Darvishi et al., 2024; Consoli et al., 2025). La evaluación de resultados de IA emerge como una competencia esencial, y los investigadores y docentes deben desarrollar habilidades para evaluar críticamente los resultados generados por IA, identificando alucinaciones, sesgos y limitaciones, y verificando la precisión y fiabilidad de la información producida (McGrath et al., 2025; Walter, 2024).

Las implicaciones para la investigación futura son igualmente sustanciales y requieren una agenda ambiciosa que aborde las brechas identificadas en esta revisión. Es necesario priorizar el estudio del impacto en el método científico, un área subrepresentada que requiere atención urgente para comprender cómo la IA está transformando el diseño metodológico, la generación de hipótesis, el análisis de datos y la difusión de resultados (Choudhary et al., 2025; Panukhnyk, 2023; Horton, 2023). Asimismo, se debe ampliar la cobertura geográfica para investigar contextos subrepresentados, especialmente en América Latina, África y otras regiones en desarrollo, y desarrollar enfoques contextualizados que consideren las necesidades y realidades específicas de estos entornos (Suárez Lima et al., 2025; Ahmad et al., 2024).

De manera complementaria, es fundamental realizar estudios longitudinales que examinen los efectos a largo plazo de la IA en el pensamiento crítico, la creatividad y la autonomía de los estudiantes e investigadores, superando las limitaciones de los diseños transversales que predominan en la literatura actual (Darvishi et al., 2024; Consoli et al., 2025). Por otro lado, se requiere investigar las alucinaciones y los sesgos para comprender cómo afectan la integridad académica y cómo se pueden mitigar mediante estrategias de verificación y formación específica (Bommasani et al., 2021; Birhane et al., 2022). Finalmente, es necesario desarrollar competencias de

prompt engineering investigando estrategias efectivas para formar a investigadores en el uso crítico de IA, de modo que puedan aprovechar al máximo estas herramientas sin caer en la dependencia o la delegación acrítica (White et al., 2023; Walter, 2024).

Las implicaciones políticas derivadas de esta revisión apuntan a la necesidad de una acción coordinada a nivel institucional, nacional e internacional para garantizar que la integración de la IA en la educación superior sea ética, equitativa y efectiva. Es imperativo desarrollar marcos regulatorios que establezcan estándares mínimos y principios rectores para el uso ético de la IA en investigación y educación, abordando cuestiones como la privacidad de datos, la transparencia algorítmica y la integridad académica (Bond et al., 2024; Mosha et al., 2026). Se debe invertir en investigación técnica y educativa para cerrar las brechas de conocimiento identificadas, financiando estudios que aborden tanto los aspectos técnicos de la IA como sus implicaciones pedagógicas y sociales (Hong et al., 2025; Sahar & Munawaroh, 2025).

Es fundamental atender la equidad asegurando que los beneficios de la IA sean accesibles para todos los estudiantes e investigadores, independientemente de su origen geográfico o socioeconómico, lo que requiere políticas de inversión en infraestructura y programas de inclusión digital (Suárez Lima et al., 2025; Ávila-Rodríguez et al., 2026). Por último, se debe promover la transparencia exigiendo que los modelos de IA utilizados en investigación y educación sean interpretables y estén documentados adecuadamente, de modo que los usuarios puedan comprender cómo funcionan, cuáles son sus limitaciones y cómo se toman las decisiones (Rudin et al., 2022).

Esta revisión, aunque diseñada y ejecutada con rigor metodológico, presenta varias limitaciones que deben ser consideradas al interpretar los hallazgos. La cobertura limitada de la búsqueda, circunscrita a bases de datos académicas indexadas y literatura gris recuperada a través de Google Scholar, implica que es posible que no se hayan incluido todos los estudios relevantes, especialmente aquellos publicados en revistas de acceso limitado o en idiomas diferentes al inglés y español (Gusenbauer & Haddaway, 2020).

La heterogeneidad metodológica de los estudios incluidos dificulta la comparación directa y la realización de metaanálisis cuantitativos, por lo que la síntesis se ha basado en un enfoque narrativo que es más susceptible a la interpretación subjetiva de los revisores (Popay et al., 2006; Higgins et al., 2019). La calidad variable de los estudios incluidos, con limitaciones en el tamaño muestral, el diseño, la validación de instrumentos y la evaluación del riesgo de sesgo, puede influir en la robustez de los hallazgos, aunque se ha realizado una evaluación sistemática de la calidad y se han priorizado los estudios de mayor calidad para las conclusiones principales (Higgins et al., 2011; Sterne et al., 2019).

La obsolescencia tecnológica es una limitación inherente a cualquier revisión sobre IA, dado que la rápida evolución de estas tecnologías significa que muchos hallazgos pueden quedar obsoletos rápidamente, y las tendencias identificadas pueden no reflejar completamente los desarrollos más recientes, especialmente en el campo de la IA generativa (McGrath et al., 2025; Walter, 2024; Bommasani et al., 2021). La concentración geográfica de los estudios en países desarrollados limita la generalización de los hallazgos a contextos con características diferentes, especialmente

en regiones en desarrollo donde la infraestructura, los recursos y los marcos regulatorios pueden ser muy distintos (Suárez Lima et al., 2025; Ahmad et al., 2024). La escasa cobertura técnica de los estudios revisados, que en su mayoría abordan la IA desde una perspectiva pedagógica sin profundizar en aspectos técnicos como arquitecturas de modelos, sesgos algorítmicos o alucinaciones, limita la profundidad del análisis desde una perspectiva estrictamente tecnológica, lo que subraya la necesidad de una mayor colaboración interdisciplinar entre expertos en IA e investigadores educativos.

A partir de los hallazgos y las limitaciones identificadas, se sugieren varias direcciones para investigaciones futuras que permitan avanzar en la comprensión del impacto de la IA en la educación superior y la investigación científica. Es necesario investigar el impacto en el método científico, examinando cómo la IA transforma el diseño metodológico, la generación de hipótesis, el análisis de datos y la difusión de resultados, y cómo estos cambios afectan la validez, fiabilidad y relevancia de la investigación (Choudhari et al., 2025; Panukhnyk, 2023; Horton, 2023).

Es crucial estudiar las alucinaciones y la fiabilidad de la IA, investigando la prevalencia y el impacto de las alucinaciones en contextos académicos y desarrollando estrategias de mitigación que permitan a investigadores y estudiantes identificar y corregir la información falsa generada por la IA (Bommasani et al., 2021; McGrath et al., 2025). Se requieren estudios longitudinales que examinen los efectos a largo plazo de la IA en el desarrollo de habilidades críticas, creatividad y autonomía, superando las limitaciones de los diseños transversales y de corto plazo que predominan en la literatura actual (Darvishi et al., 2024; Consoli et al., 2025).

Se debe ampliar la investigación en contextos subrepresentados, especialmente en América Latina, África y otras regiones en desarrollo, para comprender las necesidades, oportunidades y desafíos específicos de estos contextos y desarrollar enfoques de integración de la IA más contextualizados y culturalmente apropiados (Suárez Lima et al., 2025; Ahmad et al., 2024). Se debe evaluar la efectividad de diferentes estrategias de formación en prompt engineering, identificando los modelos más efectivos y las condiciones necesarias para su implementación exitosa en la formación de investigadores y docentes (White et al., 2023; Walter, 2024).

Es fundamental desarrollar marcos éticos y regulatorios informados por la evidencia y adaptados a diferentes contextos, que aborden cuestiones como la privacidad de datos, la equidad, la transparencia algorítmica y la integridad académica, con la participación de todos los actores relevantes (Bond et al., 2024; Khatri & Karki, 2023). Es prioridad investigar las implicaciones de la IA generativa para la evaluación, la producción de conocimiento, la propiedad intelectual y las relaciones educativas, dado que esta tecnología está transformando rápidamente el panorama educativo y la investigación (McGrath et al., 2025; Walter, 2024; Bommasani et al., 2021).

También es crucial estudiar el sesgo algorítmico y la equidad, investigando cómo los sesgos en los modelos de IA afectan la equidad en educación e investigación, y desarrollando estrategias de mitigación que aseguren que la IA no perpetúe ni amplifique las desigualdades existentes (Birhane et al., 2022; Rudin et al., 2022). Finalmente, se debe investigar la interpretabilidad y transparencia de los modelos de IA utilizados en contextos académicos, desarrollando métodos para hacer que estos

modelos sean más interpretables y transparentes, de modo que los usuarios puedan comprender cómo funcionan y confiar en sus resultados (Rudin et al., 2022).

Conclusiones

La IA tiene un impacto positivo y transformador en la educación superior, superando a los enfoques tradicionales en automatización y personalización. Sin embargo, los enfoques tradicionales mantienen valor en aspectos que requieren juicio crítico, creatividad, empatía y comprensión contextual profunda. La IA es más efectiva como complemento que potencia las capacidades humanas (inteligencia aumentada), no como sustituto de estas (Bond et al., 2024; Ruano-Borbalan, 2025; Consoli et al., 2025).

El 87.9% de los estudios se publicaron entre 2022-2026, coincidiendo con la irrupción de la IA generativa. Las principales aplicaciones son aprendizaje personalizado (62.5%), evaluación automatizada (45.3%) y apoyo a la investigación (28.1%). Los beneficios documentados incluyen eficiencia (78.1%), personalización (71.9%), democratización del conocimiento (56.3%) e innovación en investigación (43.8%). Los desafíos críticos son implicaciones éticas (84.4%), brecha digital (67.2%), falta de regulación (62.5%), erosión de habilidades (56.3%), alucinaciones de IA (31.3%) y problemas de opacidad (28.1%). La adopción docente depende de la utilidad percibida, el liderazgo institucional y las competencias digitales, incluyendo el prompt engineering (Aguado-García et al., 2025; Mosha et al., 2026; Walter, 2024; White et al., 2023).

La IA transforma el método científico: La generación de hipótesis asistida por IA, la opacidad de los modelos y los problemas de reproducibilidad requieren nuevos marcos epistemológicos para la investigación (Horton, 2023; Rudin et al., 2022). La alfabetización en IA es fundamental: Investigadores y docentes deben desarrollar competencias de prompt engineering, evaluación crítica de resultados de IA y conciencia de sesgos y alucinaciones (Walter, 2024; McGrath et al., 2025; White et al., 2023).

La equidad es un desafío central: La concentración geográfica de la producción y el acceso desigual a herramientas de IA pueden exacerbar las brechas existentes (Suárez Lima et al., 2025; Ávila-Rodríguez et al., 2026). Se necesitan marcos regulatorios: Las instituciones y los gobiernos deben desarrollar políticas claras sobre el uso ético de IA, la integridad académica y la transparencia algorítmica (Bond et al., 2024; Mosha et al., 2026).

Se recomienda a las instituciones desarrollar políticas claras, marcos éticos y programas de formación docente e investigadora, invirtiendo en equidad en el acceso y en infraestructura tecnológica (Kusumaningrum & Rosydiana, 2026; Choudhari et al., 2025). A los docentes e investigadores: Desarrollar alfabetización en IA, competencias de prompt engineering y pensamiento crítico para integrarla pedagógicamente y en la investigación (Walter, 2024; McGrath et al., 2025; White et al., 2023). A la comunidad científica, explorar el potencial de la IA para acelerar el descubrimiento, pero establecer límites claros en autoría, originalidad y transparencia metodológica (Panukhnyk, 2023; Rudin et al., 2022).

Con relación a la investigación futura, se debe priorizar el estudio del impacto de la IA en el método científico y la generación de conocimiento, ampliar cobertura geográfica,

realizar estudios longitudinales y examinar las implicaciones de la IA generativa, las alucinaciones y los sesgos algorítmicos. La IA es una oportunidad transformadora para la educación superior y la investigación científica, pero su integración debe ser guiada por principios éticos, pedagógicos y de equidad que preserven los valores fundamentales de la educación y la ciencia. La pregunta no es si la IA transformará la academia, sino cómo podemos asegurar que esta transformación sea positiva, inclusiva y sostenible (Bond et al., 2024; Ruano-Borbalan, 2025; Bommasani et al., 2021).

Acerca de

Contribución de los autores: Los autores contribuyeron a la conceptualización del estudio, al desarrollo metodológico, al análisis e interpretación de los datos, a la redacción del manuscrito y a la revisión crítica de su contenido intelectual.

Financiamiento: Los autores declaran que no recibieron financiamiento para esta investigación.

Conflicto de interés: Los autores declaran no tener conflicto de intereses.

Certificación ética: El protocolo del presente estudio fue sometido a revisión y aprobado por el Comité de Ética en Investigación de la Universidad, en cumplimiento de los principios éticos y normativas institucionales aplicables.

Historia del artículo: Artículo recibido: 28 de abril de 2026 | Arbitrado: 19 de mayo de 2026 | Aceptado: 22 de junio de 2026 | Publicado: 12 de agosto de 2026.

Referencias

- Aguado-García, J.-M., Alonso-Muñoz, S., & De-Pablos-Heredero, C. (2025). Using Artificial Intelligence for Higher Education: An Overview and Future Research Avenues. *SAGE Open*, 15(1). <https://doi.org/10.1177/21582440251340352>
- Ahmad, M., Subih, M., Fawaz, M., Alnuqaidan, H., Abuejheisheh, A., Naqshbandi, V., & Alhalaiqa, F. (2024). Awareness, benefits, threats, attitudes, and satisfaction with AI tools among Asian and African higher education staff and students. *Journal of Applied Learning & Teaching*, 7(1), 57-64. <https://doi.org/10.37074/jalt.2024.7.1.1>
- Akhmadieva, R. S., Kalmazova, N. A., Belova, T., Prokopyev, A., Molodozhnikova, N. M., & Spichak, V. (2024). Research trends in the use of artificial intelligence in higher education. *Frontiers in Education*, 9, 1438715. <https://doi.org/10.3389/feduc.2024.1438715>
- Ávila-Rodríguez, A. M., Quispe-Del-Castillo, M. A., Mendoza-León, O. E., & Herrera-Mejía, Z. E. (2026). Inteligencia artificial en la enseñanza y aprendizaje de la Educación Superior Universitaria: Una revisión sistemática. *Revista de Ciencias Sociales*, 32(1), 417-435. <https://doi.org/10.31876/rcs.v32i1.45214>
- Bond, M., Khosravi, H., de Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W., & Siemens, G. (2024). A meta systematic review of artificial intelligence in higher education: a call for increased ethics, collaboration, and rigour. *International*

- Journal of Educational Technology in Higher Education, 21(1), 4.
<https://doi.org/10.1186/s41239-023-00436-z>
- Booth, A., Sutton, A., & Papaioannou, D. (2016). *Systematic Approaches to a Successful Literature Review* (2nd ed.). SAGE Publications.
- Bramer, W. M., Giustini, D., de Jonge, G. B., Holland, L., & Bekhuis, T. (2016). De-duplication of database search results for systematic reviews in EndNote. *Journal of the Medical Library Association*, 104(3), 240-243. <https://doi.org/10.3163/1536-5050.104.3.014>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. <https://doi.org/10.1191/1478088706qp063oa>
- Castillo-Martínez, I. M., Flores-Bueno, D., Gómez-Puente, S. M., & Vite-León, V. O. (2024). AI in higher education: a systematic literature review. *Frontiers in Education*, 9, 1391485. <https://doi.org/10.3389/feduc.2024.1391485>
- Choudhari, P., Chawla, N., Lenka, R., Gopalakrishnan, G., Crasto, D., Deshmukh, D. B., & Gaidhani, S. (2025). Artificial Intelligence Reshaping Research, Innovation, and Collaboration in Higher Education: A Data-Based Analysis. *International Journal of Accounting and Economics Studies*, 13(1). <https://doi.org/10.14419/ewdxjg17>
- Consoli, T., Schmitz, M. L., Antonietti, C., Gonon, P., Cattaneo, A., & Petko, D. (2025). Quality of technology integration matters: Positive associations with students' behavioral engagement and digital competencies for learning. *Education and Information Technologies*, 30(6), 7719-7752. <https://doi.org/10.1007/s10639-024-13118-8>
- Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: the state of the field. *International Journal of Educational Technology in Higher Education*, 20(1), 22. <https://doi.org/10.1186/s41239-023-00392-8>
- Darvishi, A., Khosravi, H., Sadiq, S., Gašević, D., & Siemens, G. (2024). Impact of AI assistance on student agency. *Computers & Education*, 210, 104967. <https://doi.org/10.1016/j.compedu.2023.104967>
- Fan, O., Zheng, L., & Jiao, P. (2022). Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020. *Education and Information Technologies*, 27(6), 7893-7925. <https://doi.org/10.1007/s10639-022-10925-9>
- Figueroa de la Fuente, M., & Farhadian, G. (2025). Systematic Review of the Impact of Artificial Intelligence in Higher Education. *Journal of Teaching and Learning*, 19(4). <https://doi.org/10.22329/jtl.v19i4.9849>
- Guo, S., Zheng, Y., & Zhai, X. (2024). Artificial intelligence in education research during 2013-2023: A review based on bibliometric analysis. *Education and Information Technologies*, 29(13), 16387-16410. <https://doi.org/10.1007/s10639-024-12491-8>
- Gusenbauer, M., & Haddaway, N. R. (2020). Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. *Research Synthesis Methods*, 11(2), 181-217. <https://doi.org/10.1002/jrsm.1378>

- Harsanti, H. R., Sudibjo, N., Riady, S., & Yu, P. (2025). Exploring Indonesian teachers' intention to use artificial intelligence in schools: the impact of digital leadership, digital readiness, and perceived usefulness. *Cogent Social Sciences*, 11(1). <https://doi.org/10.1080/23311886.2025.2593596>
- Hidayati, U. (2026). Integration of Artificial Intelligence in Learning: Challenges, Ethics, and Implications for Teacher Professionalism. *Archipel: Journal of Indonesian Interdisciplinary Studies*, 1(6), 47-55. <https://doi.org/10.65739/archipel.v1i6.38>
- Higgins, J. P. T., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., ... & Sterne, J. A. C. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*, 343, d5928. <https://doi.org/10.1136/bmj.d5928>
- Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (Eds.). (2019). *Cochrane Handbook for Systematic Reviews of Interventions* (2nd ed.). Wiley-Blackwell.
- Hinojo-Lucena, F.-J., Aznar-Díaz, I., Cáceres-Reche, M.-P., & Romero-Rodríguez, J.-M. (2019). Artificial Intelligence in Higher Education: A Bibliometric Study on its Impact in the Scientific Literature. *Education Sciences*, 9(1), 51. <https://doi.org/10.3390/educsci9010051>
- Hong, T. T. M., Tung, N. T. T., & Thanh, N. T. P. (2025). Mapping artificial intelligence research in higher education toward sustainable development. *Discover Sustainability*, 6(1), 105. <https://doi.org/10.1007/s43621-025-02162-0>
- Horton, R. (2023). Artificial intelligence and the future of scientific publishing. *The Lancet*, 401(10375), 438-439. [https://doi.org/10.1016/S0140-6736\(23\)00188-9](https://doi.org/10.1016/S0140-6736(23)00188-9)
- Jabeen, G., Goli, G., & Babu, R. (2025). Transforming education with AI: Teachers' views, challenges, and pathways to success. *Journal of Engineering Education Transformations*, 38(IS2), 127-136. <https://doi.org/10.16920/jeet/2025/v38is2/25016>
- Khatri, B., & Karki, P. D. (2023). Artificial Intelligence (AI) in Higher Education: Growing Academic Integrity and Ethical Concerns. *Nepalese Journal of Development and Rural Studies*, 20(1), 69-78. <https://doi.org/10.3126/njdrs.v20i01.64134>
- Kusumaningrum, S. R., & Rosydiana, E. A. (2026). Redefining Teacher Professionalism in the Age of Artificial Intelligence: An Indonesian Perspective. *KnE Social Sciences*, 11(2), 101-109. <https://doi.org/10.18502/kss.v11i2.20731>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174. <https://doi.org/10.2307/2529310>
- Lefebvre, C., Glanville, J., Briscoe, S., Littlewood, A., Marshall, C., Metzendorf, M. I., ... & Wieland, L. S. (2022). Searching for and selecting studies. In J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, & V. A. Welch (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions* (Version 6.3). Cochrane.
- Long, H. A., French, D. P., & Brooks, J. M. (2020). Optimising the value of the critical appraisal skills programme (CASP) tool for qualitative research. *International Journal of Nursing Studies*, 110, 103713.

- Maphosa, V., & Maphosa, M. (2023). Artificial intelligence in higher education: a bibliometric analysis and topic modeling approach. *Applied Artificial Intelligence*, 37(1), 2261730. <https://doi.org/10.1080/08839514.2023.2261730>
- Marengo, A., Pagano, A., Pange, J., & Soomro, K. A. (2024). The educational value of artificial intelligence in higher education: a 10-year systematic literature review. *Interactive Technology and Smart Education*, 21(4), 571-594. <https://doi.org/10.1108/itse-11-2023-0218>
- McGrath, C., Farazouli, A., & Cerratto-Pargman, T. (2025). Generative AI chatbots in higher education: A review of an emerging research area. *Higher Education*, 89(6), 1533-1549. <https://doi.org/10.1007/s10734-024-01288-w>
- Mosha, N., Chigwada, J. P., Ketchiwou, G. F., & Ngulube, P. (2026). A Systematic Review of Artificial Intelligence in Higher Education Institutions (HEIs): Functionalities, Challenges, and Best Practices. *Education Sciences*, 16(2), 185. <https://doi.org/10.3390/educsci16020185>
- Noyes, J., Booth, A., Moore, G., Flemming, K., Tunçalp, Ö., & Shakibazadeh, E. (2018). Synthesising quantitative and qualitative evidence to inform guidelines on complex interventions: clarifying the purposes, designs and outlining some methods. *BMJ Global Health*, 4(Suppl 1), e000893.
- Ocen, S., Elasu, J., Aarakit, S., & Olupot, C. (2025). Artificial intelligence in higher education institutions: review of innovations, opportunities and challenges. *Frontiers in Education*, 10, 1530247. <https://doi.org/10.3389/feduc.2025.1530247>
- Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—a web and mobile app for systematic reviews. *Systematic Reviews*, 5(1), 210. <https://doi.org/10.1186/s13643-016-0384-4>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Panukhnyk, O. (2023). Artificial intelligence in the educational process and scientific research of higher education applicants: responsible boundaries of AI content. *Galic'kij ekonomičnij visnik*, 84(3), 202-211. https://doi.org/10.33108/galicianvisnyk_tntu2023.04.202
- Peng, J., & Li, Y. (2025). Frontiers of Artificial Intelligence for Personalized Learning in Higher Education: A Systematic Review of Leading Articles. *Applied Sciences*, 15(18), 10096. <https://doi.org/10.3390/app151810096>
- Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., ... & Duffy, S. (2006). Guidance on the conduct of narrative synthesis in systematic reviews. *ESRC Methods Programme*, 15(1), 47-71.
- Ruano-Borbalan, J.-C. (2025). The transformative impact of artificial intelligence on higher education: A critical reflection on current trends and futures directions. *International Journal of Chinese Education*, 14(1). <https://doi.org/10.1177/2212585x251319364>

- Rudin, C., Chen, C., Chen, Z., Huang, H., Semenova, L., & Zhong, C. (2022). Interpretable machine learning: Fundamental principles and 10 grand challenges. *Statistics Surveys*, 16, 1-85. <https://doi.org/10.1214/21-SS133>
- Sahar, R., & Munawaroh, M. (2025). Artificial intelligence in higher education with bibliometric and content analysis for future research agenda. *Discover Sustainability*, 6(1), 106. <https://doi.org/10.1007/s43621-025-01086-z>
- Shea, B. J., Reeves, B. C., Wells, G., Thuku, M., Hamel, C., Moran, J., ... & Henry, D. A. (2017). AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*, 358, j4008. <https://doi.org/10.1136/bmj.j4008>
- Sterne, J. A. C., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., ... & Higgins, J. P. T. (2016). ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*, 355, i4919. <https://doi.org/10.1136/bmj.i4919>
- Sterne, J. A. C., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., ... & Higgins, J. P. T. (2019). RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*, 366, l4898. <https://doi.org/10.1136/bmj.l4898>
- Suárez Lima, G. J., Ayala Velasco, E. S., Lázaro Russell, C. A., & Correa Burgos, J. A. (2025). El impacto de la inteligencia artificial en la innovación docente y la optimización del proceso de enseñanza: La evolución de la educación superior. *Revista Social Fronteriza*, 5(2), e635. [https://doi.org/10.59814/resofro.2025.5\(2\)635](https://doi.org/10.59814/resofro.2025.5(2)635)
- Walter, Y. (2024). Embracing the future of Artificial Intelligence in the classroom: The relevance of AI literacy, prompt engineering, and critical thinking in modern education. *International Journal of Educational Technology in Higher Education*, 21(1), 1-29. <https://doi.org/10.1186/s41239-024-00448-3>
- Wang, X., Jiang, X., Wei, W., Tan, S. C., Sun, D., & Zhang, Z. (2026). Navigating AI Integration Into Higher Education: A Study of the Dynamics Influencing Early-Career University Teachers' Readiness to Integrate AI in Education. *Journal of Educational Computing Research*. <https://doi.org/10.1177/07356331261456128>
- Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., ... & Fedus, W. (2022). Emergent abilities of large language models. *Transactions on Machine Learning Research*. <https://openreview.net/forum?id=yzkSU5zdWd>
- White, J., Hays, S., Fu, Q., Spencer-Smith, J., & Schmidt, D. C. (2023). A prompt pattern catalog to enhance prompt engineering with ChatGPT. *arXiv preprint arXiv:2302.11382*. <https://doi.org/10.48550/arXiv.2302.11382>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), 39. <https://doi.org/10.1186/s41239-019-0171-0>
- Zheng, J., Li, S., Wang, T., & Lajoie, S. P. (2024). Unveiling emotion dynamics in problem-solving: A comprehensive analysis with an intelligent tutoring system using facial expressions and electrodermal activities. *International Journal of Educational*

Technology in Higher Education, 21(1), 1-20. <https://doi.org/10.1186/s41239-024-00462-5>



Este artículo está bajo la licencia Creative Commons Atribución 4.0 Internacional (CC BY 4.0).