

Inteligencia artificial e investigación educativa en la innovación de la enseñanza: una revisión bibliográfica

Artificial intelligence and educational research in teaching innovation: A literature review

 Omar Astorayme Mendoza¹

 Belén Vila Ososres ²

 Mitzi Lourdes del Carmen Linares-Vizcarra ³

 Maria Magdalena Llamuca Guaman ⁴

¹Universidad César Vallejo.
Lima, Perú. Correo:
omastoraymem@ucvvirtual.edu.pe

²Universidad Científica del Sur.
Lima, Perú. Correo:
bvila@cientifica.edu.pe

³Universidad Nacional Jorge Basadre Grohmann.
Tacna, Perú. Correo:
mlinaresv@unjbg.edu.pe

⁴Universidad Autónoma de los Andes.
Ambato, Ecuador. Correo:
yquispea@unmsm.edu.pe

Cómo citar: Astorayme Mendoza, O., Vila Ososres, B., Linares-Vizcarra, M. L. del C., & Llamuca Guaman, M. M. (2026). Inteligencia artificial e investigación educativa en la innovación de la enseñanza: una revisión bibliográfica *Revista de Educación Simón Rodríguez*, 6(12), 1-18.
<https://doi.org/10.62319/simonrodriguez.v.6i12.214>

Resumen

Contexto: La inteligencia artificial se ha consolidado como un recurso emergente para transformar la enseñanza mediante personalización, tutoría, analítica del aprendizaje, evaluación automatizada y herramientas generativas. Sin embargo, su incorporación requiere fundamentos pedagógicos, éticos e institucionales. **Objetivo:** Analizar los hallazgos, tensiones y orientaciones de la investigación educativa sobre inteligencia artificial relevantes para la innovación de la enseñanza. **Metodología:** Se realizó una revisión bibliográfica sistematizada con síntesis narrativa cualitativa, orientada por los lineamientos de PRISMA 2020. La búsqueda consideró publicaciones entre 2016 y 2026 relacionadas con enseñanza, aprendizaje, evaluación, docencia, alfabetización y gobernanza educativa. De 43 registros identificados, 27 documentos integraron la síntesis cualitativa. **Resultados:** Los hallazgos evidencian que la personalización y tutoría pueden favorecer el aprendizaje cuando responden a objetivos didácticos explícitos. Asimismo, la analítica facilita la identificación de patrones, mientras que la inteligencia artificial generativa exige rediseñar la evaluación. La alfabetización digital, la formación docente y la gobernanza ética emergen como condiciones fundamentales. **Conclusión:** La innovación educativa mediante inteligencia artificial depende de su integración pedagógica, supervisión humana, evaluación empírica y marcos institucionales que garanticen equidad, privacidad, transparencia y responsabilidad.

Palabras Clave: Inteligencia artificial; Innovación educativa; Enseñanza; Alfabetización digital; Gobernanza educativa

Abstract

Context: Artificial intelligence has established itself as an emerging resource to transform teaching through personalization, tutoring, learning analytics, automated assessment, and generative tools. However, its incorporation requires pedagogical, ethical, and institutional foundations. **Objective:** To analyze the findings, tensions, and directions of educational research on artificial intelligence relevant to teaching innovation. **Methodology:** A systematized literature review with a qualitative narrative synthesis was conducted, following the PRISMA 2020 guidelines. The search considered publications between 2016 and 2026 related to teaching, learning, assessment, faculty development, literacy, and educational governance. From 43 identified records, 27 documents were included in the qualitative

synthesis. **Results:** The findings demonstrate that personalization and tutoring can favor learning when they respond to explicit didactic objectives. Likewise, analytics facilitates the identification of patterns, while generative artificial intelligence demands a redesign of assessment strategies. Digital literacy, teacher training, and ethical governance emerge as fundamental conditions. **Conclusion:** Educational innovation through artificial intelligence depends on its pedagogical integration, human oversight, empirical evaluation, and institutional frameworks that guarantee equity, privacy, transparency, and accountability.

Keywords: Artificial intelligence; Educational innovation; Teaching; Digital literacy; Educational governance

Introducción

La IA se ha convertido en un objeto inevitable de la investigación educativa por una razón doble. Por un lado, se multiplican las herramientas capaces de clasificar información, reconocer patrones, recomendar recursos, generar lenguaje, producir materiales y ofrecer respuestas conversacionales. Por otro, la incorporación de estas herramientas altera prácticas que son constitutivas de la enseñanza: planificar, explicar, formular preguntas, acompañar la práctica, retroalimentar, evaluar, documentar progresos y tomar decisiones curriculares. El interés contemporáneo no procede únicamente de la sofisticación técnica. También proviene de la posibilidad de reorganizar la relación entre datos, conocimiento, tareas y actores educativos. Por ello, estudiar IA en educación exige evitar dos reduccionismos frecuentes: la promesa de que una herramienta mejora automáticamente el aprendizaje y la idea de que todo riesgo se elimina mediante prohibiciones generales.

La literatura muestra que la IA educativa tiene una trayectoria anterior a la popularización reciente de los modelos generativos. [Zawacki-Richter et al. \(2019\)](#) identificaron, en educación superior, cuatro áreas consolidadas de aplicación: perfilado y predicción, evaluación y retroalimentación, sistemas adaptativos y personalización, y sistemas de tutoría inteligente. Esa categorización es útil porque sitúa a la IA como un conjunto heterogéneo de capacidades, no como una innovación única. Un sistema que predice riesgo de abandono, un tutor que secuenciar problemas de álgebra y un modelo generativo que ayuda a revisar un ensayo pueden compartir técnicas computacionales, pero suscitan decisiones pedagógicas, epistémicas y éticas diferentes. [Wang et al. \(2024\)](#) confirman la persistencia de estas líneas y añaden que el campo se ordena alrededor de arquitecturas y productos emergentes cuya utilidad educativa depende de su diseño, datos y contexto de uso.

La pregunta central deja de ser, entonces, si la IA debe entrar o no a las instituciones educativas. La pregunta más productiva es bajo qué condiciones contribuye a innovar la enseñanza sin comprometer la agencia de estudiantes y docentes, la justicia curricular, la privacidad o el valor formativo de la evaluación. [Bond et al. \(2024\)](#), en una metasíntesis de 66 revisiones sobre IA en educación superior, advierten que el crecimiento del área no ha ido acompañado de suficiente rigor metodológico, colaboración interdisciplinaria, contextualización ni reflexión ética. Esta advertencia resulta decisiva. Una innovación didáctica se distingue de una adopción instrumental cuando se ancla en problemas de aprendizaje claramente definidos, en teoría pedagógica, en evidencia de uso, en criterios de equidad y en mecanismos para revisar sus efectos previstos e imprevistos.

El desafío se ha intensificado con la IA generativa. A diferencia de un sistema de recomendación que opera principalmente en segundo plano, los grandes modelos de lenguaje se presentan como interlocutores capaces de explicar, resumir, traducir, proponer ejemplos, producir código o generar borradores. Su interfaz conversacional los vuelve visibles para estudiantes y docentes, y desplaza el debate hacia la autoría, la integridad académica, la calidad de la información y la evaluación. [Kasneci et al. \(2023\)](#) señalan que estos modelos pueden apoyar el aprendizaje y la enseñanza, pero también introducen riesgos relacionados con fiabilidad, sesgos, privacidad, dependencia y uso en contextos evaluativos. La rapidez de su difusión obliga a distinguir entre utilidad inmediata, evidencia educativa sostenible y legitimidad institucional.

Desde una perspectiva de investigación educativa, la IA debe estudiarse como un sistema sociotécnico. El resultado no depende únicamente del algoritmo, sino de las decisiones sobre qué datos se recogen, cómo se representan los objetivos de aprendizaje, quién interpreta las recomendaciones, qué margen de elección conserva el estudiantado y qué prácticas docentes se transforman. [Williamson y Eynon \(2020\)](#) recuerdan que la IA educativa está imbricada en historias de automatización, relaciones de poder y formas de gobernanza de datos. [Berendt et al. \(2020\)](#) añaden que la elección del aprendiz y los derechos fundamentales deben considerarse en la arquitectura misma de los sistemas. Estas aproximaciones cuestionan la neutralidad de la innovación tecnológica y desplazan el foco desde la eficiencia hacia la legitimidad pedagógica y democrática.

La evidencia sobre aplicaciones previas aconseja, además, cautela con las inferencias. En educación superior en línea, [Ouyang et al. \(2022\)](#) identificaron funciones de predicción de desempeño, recomendación de recursos, evaluación automática y mejora de la experiencia de aprendizaje. Sus hallazgos incluyen mejoras reportadas en rendimiento, participación y calidad de predicción, pero los autores también demandan integración de teorías educativas, datos de proceso en tiempo real y más investigación empírica que evalúe efectos reales. De manera convergente, [Viberg et al. \(2018\)](#) observaron que la analítica de aprendizaje presentaba gran potencial, aunque una proporción limitada de los trabajos demostraba mejora de resultados o apoyo directo a la enseñanza. La diferencia entre detectar patrones y producir cambios educativos significativos continúa siendo un problema metodológico y práctico.

Este artículo propone una revisión bibliográfica enfocada en la pregunta: ¿qué hallazgos documenta la investigación académica sobre los usos, resultados, condiciones, riesgos y líneas de acción de la IA para innovar la enseñanza? El propósito no es comparar de forma cuantitativa intervenciones heterogéneas ni declarar una eficacia uniforme de la IA. Es organizar conocimiento verificable para orientar decisiones pedagógicas e institucionales. La revisión integra trabajos publicados entre 2016 y 2025, periodo en el que se consolida la investigación sobre analítica, tutoría y personalización y, al final, emergen las revisiones sobre IA generativa, evaluación y políticas de adopción. Se presta atención a educación superior y educación escolar cuando las conclusiones permiten iluminar principios de enseñanza transferibles con las cautelas correspondientes.

La relevancia de una síntesis de este tipo es triple. En primer término, permite vincular tecnologías de IA con finalidades educativas, en vez de tratarlas como fines autónomos. En segundo término, ayuda a identificar qué resultados cuentan como evidencia: no solo precisión técnica o satisfacción declarada, sino comprensión, participación, autorregulación, calidad de la retroalimentación, equidad y sostenibilidad docente.

Finalmente, aporta un marco para la investigación educativa futura. El campo necesita diseños que combinen análisis de datos con observación de aula, estudios longitudinales, perspectivas de estudiantes y docentes, auditorías de sesgo y evaluación de consecuencias distributivas. La innovación de la enseñanza no puede reducirse a la novedad de una interfaz; debe traducirse en mejores oportunidades de aprender y enseñar

Método

El estudio se desarrolló como una revisión bibliográfica sistematizada con síntesis narrativa cualitativa, orientada a identificar, organizar y analizar la producción científica relacionada con el uso de la inteligencia artificial en la innovación de la enseñanza. Para favorecer la transparencia del proceso de identificación, selección y presentación de los estudios, se tomaron como referencia los lineamientos establecidos por PRISMA 2020 (Page et al., 2021a, 2021b).

La revisión no fue registrada prospectivamente y, por sus características, no pretende alcanzar el nivel de exhaustividad propio de una revisión sistemática basada en múltiples bases de datos especializadas. En consecuencia, se asumió como una revisión bibliográfica sistematizada, sustentada en criterios explícitos de búsqueda, selección y síntesis. Este procedimiento responde a las recomendaciones de Snyder (2019), quien señala la importancia de transparentar las decisiones adoptadas durante una revisión de literatura, y de Xiao y Watson (2019), quienes destacan la necesidad de documentar la estrategia de búsqueda, los criterios de inclusión y exclusión y el proceso de selección de los estudios.

La búsqueda bibliográfica se cerró el 18 de mayo de 2026. Se realizaron consultas temáticas en índices académicos abiertos, bases de datos académicas y páginas editoriales especializadas. Entre las principales fuentes consultadas se incluyeron Springer Nature, ScienceDirect, Taylor & Francis, Wiley, SAGE, ACM, BMJ y revistas científicas de acceso abierto. Asimismo, la búsqueda se complementó mediante rastreo retrospectivo de las referencias incluidas en revisiones recientes y estudios relevantes para el objeto de investigación.

Las estrategias de búsqueda combinaron términos en inglés y español. En inglés se emplearon expresiones como: (“artificial intelligence” OR “AI” OR “generative AI” OR ChatGPT OR “intelligent tutoring” OR “learning analytics”) AND (education OR teaching OR learning OR teacher OR assessment). En español se utilizaron combinaciones como: (“inteligencia artificial” OR “IA generativa” OR ChatGPT) AND (educación OR enseñanza OR aprendizaje OR evaluación OR profesorado). Además, se incorporaron términos asociados con analítica del aprendizaje, alfabetización en inteligencia artificial, ética, sesgo algorítmico, tutoría inteligente, evaluación automatizada, formación docente y políticas institucionales.

Los criterios de inclusión comprendieron artículos arbitrados, revisiones, metaanálisis, estudios empíricos y trabajos conceptuales publicados entre 2016 y 2026, en español o inglés, relacionados directamente con la inteligencia artificial y su aplicación en la enseñanza, el aprendizaje, la evaluación, la práctica docente, la alfabetización digital o la gobernanza educativa. Se priorizaron investigaciones cuyos resultados permitieran identificar condiciones, beneficios, riesgos o implicaciones transferibles a la práctica educativa.

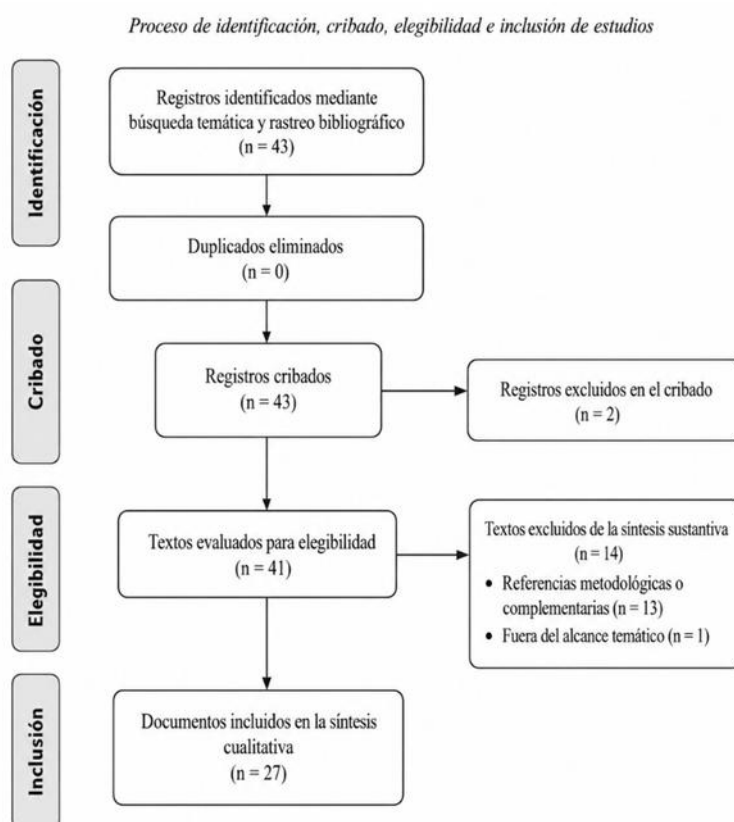
Como criterios de exclusión se consideraron documentos duplicados, trabajos centrados exclusivamente en aspectos técnicos o de ingeniería sin implicaciones educativas, comentarios o textos sin suficiente contribución analítica, publicaciones con información bibliográfica incompleta y documentos cuya temática no guardaba relación directa con los objetivos de la revisión.

La unidad de análisis estuvo constituida por cada documento seleccionado. Debido a la heterogeneidad de diseños, contextos educativos, tecnologías analizadas, poblaciones y variables de resultado, no se realizó metaanálisis. En su lugar, se optó por una síntesis narrativa cualitativa, orientada a identificar convergencias, divergencias y tendencias dentro del corpus documental.

El proceso de selección se desarrolló en tres etapas. En primer lugar, se efectuó un cribado por título y resumen para determinar la pertinencia temática. Posteriormente, se revisaron los textos seleccionados para confirmar su correspondencia con los criterios de elegibilidad. Finalmente, se analizaron los hallazgos y aportes de cada estudio en función de la pregunta de revisión. Siguiendo las recomendaciones de Peters et al. (2020), se procuró mantener coherencia entre la pregunta de investigación, los criterios de selección, el análisis de la evidencia y la presentación de los resultados.

De los 43 registros inicialmente identificados, dos fueron excluidos durante el cribado por no cumplir adecuadamente los criterios establecidos. Posteriormente, 41 documentos fueron evaluados para determinar su elegibilidad. De estos, 27 respondieron directamente a la pregunta sustantiva de la investigación y fueron incluidos en la síntesis cualitativa. Trece documentos adicionales se emplearon como soporte metodológico o para contextualizar y contrastar los resultados, mientras que un documento fue excluido por encontrarse fuera del alcance temático de la revisión.

Figura 1. *Flujo de selección de documentos*



La extracción de información se organizó en una matriz con autoría, año, tipo de estudio, contexto, aplicación de IA y hallazgos relevantes. Para la síntesis se compararon convergencias y tensiones entre cinco ejes: analítica, personalización y tutoría; evaluación; alfabetización y desarrollo profesional; ética y derechos; e IA generativa y gobernanza institucional. No se estableció una jerarquía de eficacia entre herramientas, pues ello excedería la homogeneidad del corpus. En su lugar, se buscó determinar qué condiciones acompañan a los beneficios informados y qué límites recurrentes aparecen. Esta elección permite que los resultados sean útiles para docentes, directivos e investigadores sin ocultar la pluralidad de enfoques del campo.

Resultados

El corpus final reúne predominantemente revisiones sistemáticas, revisiones de alcance, metaanálisis y trabajos conceptuales de referencia. Esta composición es adecuada para el propósito de revisar el estado de la cuestión, pero también revela una característica del área: gran parte de la literatura analiza categorías de herramientas o agrega estudios previos, mientras continúa siendo menor la disponibilidad de evaluaciones longitudinales, comparativas y situadas en el aula. [Crompton y Burke \(2023\)](#) describen una expansión rápida de la IA en educación superior y, simultáneamente, una insuficiencia de evidencias sobre usos pedagógicos, roles docentes y efectos institucionales. La revisión de [Bond et al. \(2024\)](#) llega a una conclusión compatible cuando pide más rigor, colaboración e investigación contextualizada.

En el plano de las aplicaciones, los resultados confirman una continuidad entre la IA previa a los modelos generativos y la discusión actual. La personalización, el perfilado y la tutoría inteligente siguen ocupando un lugar central; la analítica sirve para identificar trayectorias

y posibles riesgos; la evaluación automatizada procura acelerar o ampliar la retroalimentación; y la IA generativa hace visibles nuevos apoyos para diseñar y producir conocimiento. No obstante, los trabajos incluidos convergen en un punto: el efecto educativo no se deriva de la presencia de la herramienta, sino del modo en que se integra en una secuencia didáctica, de las decisiones humanas que median sus recomendaciones y de la calidad de las condiciones institucionales.

La Tabla 1 presenta los 27 documentos incluidos y sintetiza el hallazgo relevante que se utilizó en el análisis. La tabla no busca reducir los trabajos a una frase concluyente; su función es hacer visible el vínculo entre cada fuente y los temas interpretativos del artículo.

Tabla 1. Características y principales hallazgos de los estudios incluidos en la síntesis cualitativa.

N.º	Autor y año	Hallazgos relevantes
1	Zawacki-Richter et al. (2019)	La investigación de IA en educación superior se concentró en perfilado/predicción, evaluación, personalización y tutores inteligentes; participaron pocos educadores y hubo escasa reflexión ética y pedagógica.
2	Bond et al. (2024)	La metasíntesis de 66 revisiones identifica el predominio de sistemas adaptativos y personalización, y solicita más rigor metodológico, ética, contextualización e interdisciplinariedad.
3	Wang et al. (2024)	La literatura se organiza en aprendizaje adaptativo/tutoría, evaluación y gestión inteligentes, perfilado/predicción y aplicaciones emergentes; persisten vacíos teóricos y de implementación.
4	Ouyang et al. (2022)	En enseñanza superior en línea, la IA predice desempeño o satisfacción, recomienda recursos y automatiza evaluación; se requiere teoría educativa, datos de proceso y más pruebas de efectos reales.
5	Viberg et al. (2018)	La analítica de aprendizaje prometía apoyar docencia y resultados, pero la evidencia demostrada de mejora era todavía limitada.
6	Ifenthaler y Yau (2020)	La analítica puede detectar estudiantes en riesgo y apoyar continuidad, aunque faltan estudios rigurosos, de gran escala y con evidencia de eficacia sostenida.
7	Larrabee Sønderlund et al. (2019)	Una proporción pequeña de estudios evaluó intervenciones de analítica; la calidad metodológica no permitía establecer eficacia ni generalización concluyentes.
8	Mousavinasab et al. (2021)	Los tutores inteligentes integran modelos de dominio y estudiante, reglas, minería de datos y redes bayesianas; son insuficientes la evaluación experimental y la implementación móvil.
9	Kulik y Fletcher (2016)	El metaanálisis de 50 evaluaciones controladas informó un efecto medio favorable de los tutores inteligentes frente a enseñanza convencional, condicionado por evaluación e implementación.
10	Kabudi et al. (2021)	Los sistemas adaptativos personalizan rutas con diversos datos y técnicas, pero aún son escasas las soluciones plenamente implementadas para problemas educativos reales.

N.º	Autor y año	Hallazgos relevantes
11	Ramesh y Sanampudi (2022)	La puntuación automática de ensayos tiene dificultades para captar contenido, cohesión y coherencia; la revisión humana y criterios transparentes son indispensables.
12	Long y Magerko (2020)	La alfabetización en IA comprende comprender, usar, evaluar críticamente y crear con IA, bajo principios de diseño centrados en el aprendiz.
13	Ng et al. (2021)	Propone cuatro dimensiones de alfabetización: conocer/comprender, usar/aplicar, evaluar/crear y ética; estas orientan currículo y evaluación.
14	Holmes et al. (2022)	Un marco ético para IA educativa debe integrar justicia, responsabilidad, transparencia, sesgo, autonomía, agencia e inclusión.
15	Baker y Hawn (2022)	El sesgo aparece a lo largo de todo el ciclo de datos y algoritmos; la equidad exige auditorías, representatividad y supervisión pedagógica sostenida.
16	Akgun y Greenhow (2022)	En la educación escolar, privacidad, vigilancia, sesgo y decisiones automatizadas deben abordarse con alfabetización en IA y criterio ético docente.
17	Williamson y Eynon (2020)	Los efectos de la IA dependen de relaciones históricas de poder, datos y gobernanza; no se explican únicamente por la capacidad técnica.
18	Kasneci et al. (2023)	Los modelos de lenguaje pueden apoyar a estudiantes y docentes, pero introducen problemas de fiabilidad, sesgo, privacidad, dependencia y evaluación, por lo que requieren supervisión humana.
19	Chiu et al. (2023)	La IA adopta múltiples roles en aprendizaje, enseñanza, evaluación y administración, y enfrenta desafíos técnicos, pedagógicos, éticos e institucionales.
20	Lo (2023)	ChatGPT puede funcionar como tutor y apoyo docente, pero su rendimiento es desigual entre disciplinas y obliga a rediseñar los métodos de evaluación.
21	Tlili et al. (2023)	El caso ChatGPT revela oportunidades de apoyo junto con riesgos de engaño, información errónea, privacidad y sesgo que exigen adopción responsable.
22	Cotton et al. (2023)	La integridad académica demanda políticas claras, formación, rediseño de la evaluación y modalidades diversas; la detección por sí sola es insuficiente.
23	Xia et al. (2024)	La evaluación debe transformarse para desarrollar autorregulación, aprendizaje responsable e integridad, con desarrollo profesional docente y revisión de políticas.
24	Tan et al. (2025)	Existe una brecha entre la formación docente ofrecida y las necesidades prácticas para integrar IA; se necesitan programas situados, continuos y pertinentes.
25	Crompton y Burke (2023)	El crecimiento del campo convive con falta de investigación sobre usos pedagógicos, docencia y efectos institucionales de la IA.
26	Sperling et al. (2024)	La alfabetización en IA está insuficientemente representada en la formación docente, especialmente en sus dimensiones prácticas y éticas.
27	Jin et al. (2025)	La adopción universitaria de IA generativa requiere orientaciones institucionales explícitas y gobernanza capaz de acompañar la diversidad de contextos.

El primer hallazgo transversal es que la IA ofrece mayor potencial pedagógico cuando se concibe como un recurso para aumentar la capacidad de observar, interpretar y responder a la diversidad del aprendizaje. Los sistemas adaptativos y de tutoría constituyen el ejemplo más estudiado. [Kulik y Fletcher \(2016\)](#), en un metaanálisis de 50 evaluaciones, hallaron un efecto medio favorable de los sistemas tutores inteligentes frente a la enseñanza convencional. Este resultado no autoriza una conclusión universal: los efectos cambiaron según la prueba utilizada, el tratamiento de comparación y las condiciones de implementación. Más que prometer una sustitución de la docencia, el metaanálisis sugiere que la tutoría inteligente puede ser útil para ofrecer práctica guiada y retroalimentación específica cuando se diseña coherentemente con los objetivos y el contexto.

[Mousavinasab et al. \(2021\)](#) muestran que los tutores inteligentes combinan modelos del dominio, modelos del estudiante y métodos como reglas, minería de datos o redes bayesianas. Su aporte principal es recordar que la adaptación es una hipótesis de diseño: el sistema decide qué necesita el estudiante a partir de una representación parcial del conocimiento y de la actividad. Si esa representación es pobre, opaca o descontextualizada, la personalización puede convertirse en clasificación rígida. [Kabudi et al. \(2021\)](#) encuentran, además, que muchos sistemas adaptativos siguen siendo prototipos o no se implementan plenamente frente a problemas educativos reales. La distancia entre la prueba técnica y la práctica docente explica por qué la innovación requiere acompañamiento, infraestructura, formación y evaluación continuada.

En la analítica de aprendizaje, la posibilidad de anticipar dificultades es atractiva, en especial en entornos virtuales con grandes cantidades de trazas de interacción. [Ifenthaler y Yau \(2020\)](#) concluyen que las técnicas analíticas pueden apoyar el éxito académico y detectar estudiantes en riesgo, pero advierten la escasez de evidencias rigurosas y a gran escala. [Larrabee Sønderlund et al. \(2019\)](#) son aún más cautos: de una amplia literatura, solo un número reducido de estudios evaluaba intervenciones, y la base disponible no era suficiente para afirmar eficacia o generalización sólidas. [Stojanov y Daniel \(2024\)](#) amplían el panorama al mostrar que el uso de big data y analítica suele orientarse al apoyo del aprendizaje y, en menor medida, a la docencia y la administración, incorporando también preocupaciones de bienestar.

Esta evidencia orienta una distinción crucial para la innovación de la enseñanza. Un tablero que identifica bajo acceso a materiales puede ser útil para iniciar una conversación, pero no debe confundirse con una explicación de por qué un estudiante aprende menos ni con una decisión pedagógica automática. La interpretación debe considerar condiciones de conectividad, trabajo, discapacidad, idioma, experiencia previa, diseño de la tarea y participación en múltiples espacios. La mejor función de la IA analítica no es reemplazar el juicio profesional, sino proporcionar indicios que el docente pueda contrastar con conocimiento situado, diálogo y evidencia cualitativa.

El segundo hallazgo transversal corresponde a la evaluación. La puntuación automatizada puede agilizar procesos y generar retroalimentación inmediata, pero [Ramesh y Sanampudi \(2022\)](#) evidencian límites para capturar relevancia de contenido, coherencia y cohesión. Tales dimensiones son centrales en el aprendizaje de escritura argumentativa y no pueden delegarse sin especificar qué se valora y quién responde cuando la herramienta se equivoca. La evaluación automatizada es más defendible cuando se utiliza como apoyo de bajo

riesgo, con rúbricas públicas, muestras de contraste, revisión humana y posibilidades de apelación. No basta con medir concordancia con calificaciones humanas; se debe examinar si la retroalimentación mejora las revisiones del estudiante y si distribuye errores de manera desigual entre grupos.

La aparición de IA generativa vuelve esta cuestión más visible y prepara el tránsito hacia el análisis de sus implicaciones. [Xia et al. \(2024\)](#) proponen reorientar la evaluación para fortalecer autorregulación, aprendizaje responsable e integridad. Esta propuesta no equivale a renunciar a toda comprobación individual del aprendizaje. Implica, más bien, diversificar evidencias: procesos de elaboración, justificaciones orales, bitácoras, defensas, revisión entre pares, tareas aplicadas a contextos locales y reflexión documentada sobre el uso de herramientas. Así, la IA puede ser un objeto de análisis y una ayuda declarada, en lugar de permanecer como una presencia clandestina que obliga a endurecer la vigilancia.

Discusión

Los resultados permiten sostener que la IA puede innovar la enseñanza, pero solo cuando se articula con una teoría de la acción pedagógica. Esta condición modifica el criterio de adopción. En lugar de preguntar qué herramienta es más novedosa o qué plataforma automatiza más tareas, conviene iniciar con una pregunta de enseñanza: ¿qué obstáculo de comprensión, práctica, participación, retroalimentación y organización se pretende abordar? La respuesta delimita qué datos son pertinentes, qué intervención es proporcional y cómo se podrá valorar su efecto. Sin esta secuencia, la personalización corre el riesgo de convertirse en una secuencia automatizada de contenidos; la analítica, en vigilancia de actividad; y la IA generativa, en producción rápida de textos sin construcción de conocimiento.

La literatura sobre aprendizaje adaptativo ilustra esta necesidad. [Fontaine et al. \(2019\)](#), en una revisión y metaanálisis sobre e-learning adaptativo para profesionales y estudiantes de salud, hallaron efectos favorables en habilidades, aunque la evidencia de conocimiento fue menos concluyente y coexistieron heterogeneidad y riesgo de sesgo. Más recientemente, [du Plooy et al. \(2024\)](#) reportan que una proporción relevante de los estudios sobre aprendizaje adaptativo personalizado informa mejoras de rendimiento y, en menor medida, de compromiso; no obstante, subrayan la diversidad de plataformas, indicadores y estrategias de implementación. La conclusión prudente no es que toda adaptación funcione, sino que la adaptación necesita definir con precisión qué se ajusta, dificultad, ritmo, ejemplo, modalidad, feedback o trayectoria, con qué evidencia y para qué resultado educativo.

La investigación sobre sistemas tutores añade un criterio de diseño: la inteligencia del sistema debe hacer visible y mejorable el modelo pedagógico que contiene. El metaanálisis de [Kulik y Fletcher \(2016\)](#) ofrece evidencia alentadora, mientras que [Mousavinasab et al. \(2021\)](#) y [Kabudi et al. \(2021\)](#) advierten variación técnica, evaluación experimental incompleta y baja transferencia a contextos reales. [Létourneau et al. \(2025\)](#) encuentran en K-12 resultados generalmente positivos de tutores inteligentes, aunque atenuados frente a sistemas no inteligentes y dependientes de duración, muestras y contexto. Para los equipos docentes, esto implica que un tutor no debe adoptar una secuencia cerrada sin revisión

curricular. Sus explicaciones, ejercicios, criterios de progreso y formas de error deben dialogar con los objetivos de la asignatura, las características del grupo y la posibilidad de que el estudiante explique, cuestionen y use la retroalimentación.

La personalización merece una discusión especial porque suele presentarse como el beneficio más evidente de la IA. Personalizar puede significar ofrecer ejemplos próximos a los intereses del estudiante, adecuar el andamiaje, proponer práctica adicional o recomendar recursos en función de una necesidad diagnosticada. Estas operaciones pueden ampliar oportunidades si se mantienen reversibles y comprensibles. Sin embargo, un modelo que predice bajo desempeño puede bajar expectativas, restringir acceso a tareas complejas o etiquetar a un estudiante con categorías difíciles de cuestionar. La innovación responsable adopta la predicción como una hipótesis revisable y no como sentencia. Debe ofrecer explicaciones legibles, preservar el derecho a disentir, permitir corrección de datos y evaluar si las recomendaciones benefician por igual a grupos con trayectorias diferentes.

Aquí adquiere relevancia la aportación de [Baker y Hawn \(2022\)](#). El sesgo algorítmico no es un incidente aislado que se corrige al final del desarrollo, sino un riesgo que puede originarse en la definición de la variable, la selección de la muestra, el etiquetado, el diseño del modelo, el modo de implementación y la interpretación de resultados. En educación, donde las desigualdades previas afectan la disponibilidad y la calidad de los datos, una métrica aparentemente neutral puede reproducir exclusiones. [Holmes et al. \(2022\)](#) sitúan en el centro la justicia, la responsabilidad, la transparencia, la autonomía, la agencia, la inclusión y el sesgo. Por ello, cualquier innovación que use datos de estudiantes debe incorporar evaluación de impacto distributivo, participación de los afectados y un proceso claro para impugnar decisiones automatizadas.

La privacidad también debe tratarse como condición pedagógica, y no como mero trámite normativo. Datos de clics, tiempo de pantalla, respuestas, interacciones conversacionales o rasgos inferidos pueden permitir personalizar apoyos, pero también ampliar la vigilancia. [Akgun y Greenhow \(2022\)](#) muestran que, en K-12, los desafíos comprenden privacidad, sesgo, vigilancia y decisiones automatizadas. De manera complementaria, [Gouseti et al. \(2024\)](#) sintetizan que privacidad, transparencia, responsabilidad y equidad deben formar parte de las políticas y currículos escolares. [Berendt et al. \(2020\)](#) colocan la elección del aprendiz y los derechos fundamentales como lentes indispensables. Una política de adopción responsable debe definir minimización de datos, propósito explícito, retención limitada, controles de acceso, seguridad, información comprensible y alternativas razonables para quienes no consientan determinados tratamientos. Tales medidas protegen derechos y, además, favorecen la confianza necesaria para que docentes y estudiantes utilicen la tecnología de modo reflexivo.

La alfabetización en IA emerge, por consiguiente, como una competencia institucional y no solo individual. [Long y Magerko \(2020\)](#) proponen que una persona alfabetizada en IA pueda reconocer sus principios y usos, evaluar críticamente sus resultados y participar en actividades de creación o diseño. [Ng et al. \(2021\)](#) organizan esta alfabetización en conocer y comprender, usar y aplicar, evaluar y crear, e incorporar cuestiones éticas. Estos marcos ayudan a superar dos extremos: la capacitación exclusivamente instrumental, que enseña a formular instrucciones sin analizar resultados, y el rechazo abstracto, que critica la IA sin

comprender cómo opera. La alfabetización de estudiantes debe incorporar verificación de fuentes, detección de alucinaciones, atribución, privacidad, sesgo, trazabilidad y reflexión sobre cuándo no usar una herramienta. La alfabetización docente necesita, además, vincular estas dimensiones con currículo, evaluación, inclusión y gestión de aula.

La brecha de formación documentada por [Sperling et al. \(2024\)](#) y [Tan et al. \(2025\)](#) es especialmente relevante. La primera revisión identifica presencia insuficiente de alfabetización en IA en formación de docentes, con subrepresentación de dimensiones prácticas y éticas. La segunda encuentra distancia entre las oportunidades de desarrollo profesional y las necesidades prácticas de quienes enseñan. Esto invalida estrategias institucionales que se limiten a proporcionar licencias de herramientas o guías generales. El desarrollo profesional debería ser sostenido, disciplinar y basado en problemas reales de docencia: diseñar una actividad con uso declarado de IA, revisar una rúbrica, analizar respuestas generadas, contrastar la retroalimentación automatizada con criterios de calidad, explorar riesgos de datos o preparar una explicación para estudiantes. Asimismo, necesita tiempo reconocido y comunidades de práctica que permitan compartir dilemas y evidencias.

La IA generativa intensifica estas exigencias porque es capaz de producir artefactos lingüísticos plausibles con un costo muy bajo. [Chiu et al. \(2023\)](#) identifican roles de la IA en aprendizaje, enseñanza, evaluación y administración, junto con desafíos interrelacionados. [Lo \(2023\)](#) concluye que ChatGPT muestra posibilidades como tutor y apoyo docente, pero ofrece desempeño desigual por disciplinas y amenaza la validez de métodos de evaluación basados solo en productos textuales. [Tili et al. \(2023\)](#) documentan simultáneamente entusiasmo por las oportunidades y preocupaciones por engaño, veracidad, privacidad y sesgo. En conjunto, estos estudios desplazan el centro de la innovación desde «detectar IA» hacia «rediseñar para aprender con y acerca de IA».

El rediseño evaluativo debe distinguir entre tareas que buscan comprobar dominio individual y tareas que pretenden desarrollar desempeño profesional o colaborativo. En el primer caso, pueden requerirse evidencias sin asistencia, instancias orales, resolución situada o explicaciones de decisiones. En el segundo, puede ser formativo que el estudiante use IA para producir una primera versión, contraste fuentes, documente indicaciones, critique sesgos y justifique cambios. [Cotton et al. \(2023\)](#) recomiendan una respuesta institucional que combine políticas, formación y rediseño de la evaluación, en lugar de depender exclusivamente de mecanismos de detección. [Albadarin et al. \(2024\)](#), en su revisión de estudios empíricos sobre ChatGPT, reportan usos para explicación, escritura, retroalimentación y aprendizaje personalizado, pero advierten que un uso excesivo puede perjudicar capacidades innovadoras y colaboración. La consigna pedagógica no debería ser usar IA en toda tarea, sino decidir cuándo el uso añade valor y qué operaciones intelectuales deben mantenerse explícitas y atribuibles al estudiante.

[Xia et al. \(2024\)](#) ofrecen una vía concreta al proponer que la transformación de la evaluación fortalezca la autorregulación, el aprendizaje responsable y la integridad. Esto implica valorar el proceso, no solo el resultado: especificar los propósitos del uso de IA, conservar registros de interacción cuando sean pertinentes, comparar versiones, verificar afirmaciones y explicar qué se aceptó o rechazó de una respuesta generada. El desarrollo

de estas prácticas puede mejorar la evaluación aun cuando no se use IA, pues favorece metacognición, transparencia de criterios y retroalimentación de calidad. La integridad académica se transforma así de un problema policial a una oportunidad para enseñar atribución, responsabilidad y juicio epistémico.

La dimensión institucional es tan importante como la didáctica. Jin et al. (2025) analizan políticas de adopción de IA generativa en universidades de diversas regiones y muestran la necesidad de orientación y gobernanza explícitas. An et al. (2025), al examinar directrices de instituciones de educación superior, identifican temas recurrentes de orientación a docentes, políticas de curso, integridad, privacidad y ética. Estas evidencias sugieren que las instituciones deben elaborar marcos vivos, revisables y participativos, no reglamentos reactivos. Tales marcos deberían delimitar usos permitidos y no permitidos por actividad, responsabilidades del proveedor y de la institución, normas de atribución, protección de datos, apoyo a la accesibilidad, procedimientos de revisión y espacios de consulta pedagógica.

Los resultados también exponen límites relevantes de la evidencia. Primero, el predominio de revisiones y estudios de alcance dificulta inferir causalidad o comparar efectos entre tecnologías. Segundo, las categorías «IA», «personalización» o «IA generativa» abarcan diseños muy distintos, por lo que los resultados no deben generalizarse indiscriminadamente. Tercero, buena parte de la investigación se concentra en educación superior, disciplinas STEM y contextos con infraestructura digital, lo que restringe su transferibilidad. Cuarto, los indicadores de éxito suelen privilegiar precisión, calificaciones, satisfacción o participación, mientras son menos frecuentes medidas de comprensión profunda, emancipación, bienestar, relación pedagógica y justicia distributiva. Quinto, el rápido cambio de modelos y políticas puede volver obsoletas algunas conclusiones técnicas; por ello, los principios pedagógicos y de derechos adquieren valor estabilizador.

Este manuscrito presenta algunas limitaciones. Se desarrolló como una revisión bibliográfica sistematizada y no como una revisión sistemática registrada; además, la búsqueda se apoyó en servicios de descubrimiento académico y fuentes abiertas, sin reproducirse de forma exhaustiva en todas las bases de datos de suscripción. La selección del corpus se realizó con base en la pertinencia temática y la disponibilidad de información bibliográfica suficiente, sin aplicar doble cribado independiente ni una evaluación formal del riesgo de sesgo de cada estudio. Asimismo, aunque se incluyeron publicaciones recientes y la búsqueda se cerró en mayo de 2026, el rápido desarrollo de la inteligencia artificial generativa hace necesaria la actualización periódica de la evidencia. Estas limitaciones permiten delimitar el alcance de las conclusiones y deben considerarse al interpretar los resultados de la síntesis.

Conclusiones

La revisión indica que la IA puede contribuir a la innovación de la enseñanza cuando se utiliza para ampliar, y no sustituir, la capacidad profesional de diagnosticar necesidades, ofrecer práctica y retroalimentación, diversificar recursos y diseñar evaluaciones más

reflexivas. Los sistemas adaptativos, la tutoría inteligente y la analítica muestran potencial, pero sus beneficios dependen de modelos pedagógicos explícitos, evidencias de implementación y revisión humana. La IA generativa acentúa la urgencia de estas condiciones porque altera las prácticas de producción textual, explicación y evaluación de manera inmediata.

La principal implicación es que la innovación pedagógica no se mide por la presencia de IA, sino por la calidad de los cambios que introduce en las oportunidades de aprender, participar y demostrar comprensión. Una adopción responsable requiere cinco compromisos: definir problemas de enseñanza antes de elegir herramientas; mantener al docente y al estudiante como agentes interpretativos; integrar alfabetización crítica en IA en currículo y formación; proteger derechos mediante privacidad, equidad, transparencia y rendición de cuentas; y evaluar efectos en condiciones reales, con atención a grupos y contextos diversos. Cuando estos compromisos faltan, la automatización puede reforzar desigualdades o debilitar la relación pedagógica. Cuando se cumplen, la IA puede convertirse en un recurso para una enseñanza más situada, dialogante, inclusiva y basada en evidencia.

La agenda de investigación debería avanzar desde demostraciones de capacidad hacia estudios que documenten procesos y consecuencias. Son necesarias investigaciones longitudinales y multimétodo, comparaciones transparentes entre diseños didácticos, análisis de impacto distributivo, estudios de carga docente, evaluación de retroalimentación generada y participación activa de estudiantes, docentes y comunidades en el diseño y la gobernanza. De este modo, la investigación educativa podrá distinguir la innovación duradera de la adopción apresurada y contribuir a que la IA sirva a finalidades educativas públicamente deliberadas.

Acerca de

Contribución de los autores: Los autores contribuyeron a la conceptualización del estudio, desarrollo metodológico, análisis e interpretación de los datos, redacción del manuscrito y 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: Por tratarse de una revisión bibliográfica basada en publicaciones científicas y fuentes documentales, el estudio no involucró participantes humanos ni requirió consentimiento informado o aprobación de un comité de ética.

Historia del artículo: Artículo recibido: 03 de junio 2026 | Arbitrado: 22 de junio 2026 | Aceptado: 15 de julio 2026 | Publicado 28 de agosto 2026.

Referencias

Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics*, 2(3), 431–440. <https://doi.org/10.1007/s43681-021-00096-7>

- Albadarin, Y., Saqr, M., Pope, N., & Tukiainen, M. (2024). A systematic literature review of empirical research on ChatGPT in education. *Discover Education*, 3. <https://doi.org/10.1007/s44217-024-00138-2>
- An, Y., Yu, J. H., & James, S. (2025). Investigating the higher education institutions' guidelines and policies regarding the use of generative AI in teaching, learning, research, and administration. *International Journal of Educational Technology in Higher Education*, 22(1). <https://doi.org/10.1186/s41239-025-00507-3>
- Baker, R. S., & Hawn, A. (2022). Algorithmic bias in education. *International Journal of Artificial Intelligence in Education*, 32(4), 1052–1092. <https://doi.org/10.1007/s40593-021-00285-9>
- Berendt, B., Littlejohn, A., & Blakemore, M. (2020). AI in education: Learner choice and fundamental rights. *Learning, Media and Technology*, 45(3), 312–324. <https://doi.org/10.1080/17439884.2020.1786399>
- 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). <https://doi.org/10.1186/s41239-023-00436-z>
- Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 4, 100118. <https://doi.org/10.1016/j.caeai.2022.100118>
- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228–239. <https://doi.org/10.1080/14703297.2023.2190148>
- 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). <https://doi.org/10.1186/s41239-023-00392-8>
- du Plooy, E., Casteleijn, D., & Franzsen, D. (2024). Personalized adaptive learning in higher education: A scoping review of key characteristics and impact on academic performance and engagement. *Heliyon*, 10(21), e39630. <https://doi.org/10.1016/j.heliyon.2024.e39630>
- Fontaine, G., Cossette, S., Maheu-Cadotte, M.-A., Mailhot, T., Deschênes, M.-F., Mathieu-Dupuis, G., Côté, J., Gagnon, M.-P., & Dubé, V. (2019). Efficacy of adaptive e-learning for health professionals and students: A systematic review and meta-analysis. *BMJ Open*, 9(8), e025252. <https://doi.org/10.1136/bmjopen-2018-025252>
- Gouseti, A., James, F., Fallin, L., & Burden, K. (2024). The ethics of using AI in K-12 education: A systematic literature review. *Technology, Pedagogy and Education*, 34(2), 161–182. <https://doi.org/10.1080/1475939X.2024.2428601>
- Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., Santos, O. C., Rodrigo, M. M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2022). Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education*, 32(3), 504–526. <https://doi.org/10.1007/s40593-021-00239-1>
- Ifenthaler, D., & Yau, J. Y.-K. (2020). Utilising learning analytics to support study success in higher education: A systematic review. *Educational Technology Research and Development*, 68(4), 1961–1990. <https://doi.org/10.1007/s11423-020-09788-z>

- Jin, Y., Yan, L., Echeverria, V., Gašević, D., & Martinez-Maldonado, R. (2025). Generative AI in higher education: A global perspective of institutional adoption policies and guidelines. *Computers and Education: Artificial Intelligence*, 8, 100348. <https://doi.org/10.1016/j.caeai.2024.100348>
- Kabudi, T., Pappas, I., & Olsen, D. H. (2021). AI-enabled adaptive learning systems: A systematic mapping of the literature. *Computers and Education: Artificial Intelligence*, 2, 100017. <https://doi.org/10.1016/j.caeai.2021.100017>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeiffer, F., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., et al (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of intelligent tutoring systems: A meta-analytic review. *Review of Educational Research*, 86(1), 42–78. <https://doi.org/10.3102/0034654315581420>
- Larrabee Sønderlund, A., Hughes, E., & Smith, J. (2019). The efficacy of learning analytics interventions in higher education: A systematic review. *British Journal of Educational Technology*, 50(5), 2594–2618. <https://doi.org/10.1111/bjet.12720>
- Létourneau, A., Deslandes Martineau, M., Charland, P., Karran, J. A., Boasen, J., & Léger, P. M. (2025). A systematic review of AI-driven intelligent tutoring systems (ITS) in K-12 education. *npj Science of Learning*, 10(1). <https://doi.org/10.1038/s41539-025-00320-7>
- Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, 13(4), 410. <https://doi.org/10.3390/educsci13040410>
- Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). Association for Computing Machinery. <https://doi.org/10.1145/3313831.3376727>
- Mousavinasab, E., Zarifsanaiey, N., R. Niakan Kalhori, S., Rakhshan, M., Keikha, L., & Ghazi Saeedi, M. (2021). Intelligent tutoring systems: A systematic review of characteristics, applications, and evaluation methods. *Interactive Learning Environments*, 29(1), 142–163. <https://doi.org/10.1080/10494820.2018.1558257>
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, 100041. <https://doi.org/10.1016/j.caeai.2021.100041>
- Ouyang, F., 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>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021a). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>

- Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... McKenzie, J. E. (2021b). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ*, 372, n160. <https://doi.org/10.1136/bmj.n160>
- Peters, M. D. J., Marnie, C., Tricco, A. C., Pollock, D., Munn, Z., Alexander, L., McInerney, P., Godfrey, C. M., & Khalil, H. (2020). Updated methodological guidance for the conduct of scoping reviews. *JBIE Evidence Synthesis*, 18(10), 2119–2126. <https://doi.org/10.11124/JBIES-20-00167>
- Ramesh, D., & Sanampudi, S. K. (2022). An automated essay scoring systems: A systematic literature review. *Artificial Intelligence Review*, 55(3), 2495–2527. <https://doi.org/10.1007/s10462-021-10068-2>
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339. <https://doi.org/10.1016/j.jbusres.2019.07.039>
- Sperling, K., Stenberg, C.-J., McGrath, C., Åkerfeldt, A., Heintz, F., & Stenliden, L. (2024). In search of artificial intelligence (AI) literacy in teacher education: A scoping review. *Computers and Education Open*, 6, 100169. <https://doi.org/10.1016/j.caeo.2024.100169>
- Stojanov, A., & Daniel, B. K. (2024). A decade of research into the application of big data and analytics in higher education: A systematic review of the literature. *Education and Information Technologies*, 29(5), 5807–5831. <https://doi.org/10.1007/s10639-023-12033-8>
- Tan, X., Cheng, G., & Ling, M. H. (2025). Artificial intelligence in teaching and teacher professional development: A systematic review. *Computers and Education: Artificial Intelligence*, 8, 100355. <https://doi.org/10.1016/j.caeai.2024.100355>
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10(1). <https://doi.org/10.1186/s40561-023-00237-x>
- Viberg, O., Hatakka, M., Bälter, O., & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. *Computers in Human Behavior*, 89, 98–110. <https://doi.org/10.1016/j.chb.2018.07.027>
- Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. *Expert Systems with Applications*, 252, 124167. <https://doi.org/10.1016/j.eswa.2024.124167>
- Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology*, 45(3), 223–235. <https://doi.org/10.1080/17439884.2020.1798995>
- Xia, Q., Weng, X., Ouyang, F., Lin, T. J., & Chiu, T. K. F. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education*, 21(1). <https://doi.org/10.1186/s41239-024-00468-z>
- Xiao, Y., & Watson, M. (2019). Guidance on conducting a systematic literature review. *Journal of Planning Education and Research*, 39(1), 93–112. <https://doi.org/10.1177/0739456X17723971>

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). <https://doi.org/10.1186/s41239-019-0171-0>