

# La inteligencia artificial en educación: consideraciones éticas y fomento al pensamiento crítico

Artificial intelligence in education: Ethical considerations and the promotion of critical thinking

MARÍA JOSÉ NOZATO LÓPEZ

## María José Nozato López.

Tecnológico Nacional de México - Instituto Tecnológico de Agua Prieta, Sonora, México. Es Ingeniera Química con especialidad en Industrial y Maestra en Ingeniería Administrativa. Actualmente cursa el Doctorado en Administración e Innovación Empresarial. Cuenta con experiencia docente en los niveles de secundaria, educación media superior y licenciatura, desde el año 2014. Como directora de educación media superior ha liderado programas académicos orientados a mejorar el rendimiento estudiantil y optimizar procesos educativos. Correo electrónico: mjnozato@gmail.com. ORCID: <https://orcid.org/0009-0003-5677-374X>.

## Resumen

Este estudio realiza una revisión sistemática de la literatura sobre el uso de la inteligencia artificial (IA) en educación, enfocándose en las consideraciones éticas y el fomento del pensamiento crítico. Ante el creciente uso de herramientas de IA en la educación, es fundamental reflexionar sobre los desafíos y oportunidades. La UNESCO ha establecido la necesidad de principios éticos para garantizar que estas tecnologías respeten los derechos humanos y contribuyan al bienestar. El análisis aborda temas como equidad, privacidad, transparencia y el papel de la IA en desarrollar el pensamiento crítico, aspectos identificados en informes de la UNESCO sobre el impacto de la IA en la educación. A través de la revisión de investigaciones, se identifican buenas prácticas y recomendaciones para que las instituciones implementen códigos de ética y políticas que aseguren el uso responsable de la IA, priorizando el bienestar de los estudiantes y el fortalecimiento de sus capacidades de análisis y cuestionamiento. Esto se alinea con el énfasis de la UNESCO en el papel clave de los educadores para fomentar el pensamiento crítico ante las tecnologías emergentes. La implementación de la inteligencia artificial en la educación no solo requiere el acceso a estas herramientas, sino también la capacidad de los docentes para guiar a los estudiantes en su uso reflexivo y ético. Así, los educadores juegan un rol esencial al promover un enfoque crítico que permita a los estudiantes analizar y cuestionar el impacto de estas tecnologías en su aprendizaje y en la sociedad.

*Palabras clave:* Inteligencia artificial, educación, ética, pensamiento crítico.

## Abstract

This study conducts a systematic literature review on the use of artificial intelligence (AI) in education, focusing on ethical considerations and the promotion of critical thinking. Given the increasing use of AI tools in education, it is essential to reflect on the associated challenges and opportunities. UNESCO has established the need for ethical principles to ensure that these technologies respect human rights and contribute to well-being. The analysis addresses topics such as equity, privacy, transparency, and the role of AI in fostering critical thinking, as identified in UNESCO reports on the impact of AI on education. Through the review of existing research, best practices and recommendations are identified for institutions to implement ethical codes and policies that ensure the responsible use of AI, prioritizing stu-

dents' well-being and enhancing their analytical and questioning abilities. This aligns with UNESCO's emphasis on the key role of educators in fostering critical thinking in the face of emerging technologies. The implementation of artificial intelligence in education requires not only access to these tools but also the capability of educators to guide students in their reflective and ethical use. Thus, educators play an essential role in promoting a critical approach that allows students to analyze and question the impact of these technologies on their learning and society.

*Keywords:* Artificial intelligence, education, ethics, critical thinking.

---

## INTRODUCCIÓN

La revisión sistemática de la literatura ha puesto de relieve el impacto creciente de la inteligencia artificial –IA– en la educación a nivel superior, enfocándose particularmente en sus implicaciones éticas y su potencial para promover el pensamiento crítico entre los estudiantes. Este estudio ha identificado desafíos éticos significativos, como los sesgos algorítmicos en sistemas de IA, que pueden amplificar desigualdades y generar discriminación (Popenici y Kerr, 2017; Williamson, 2019). Además, se ha subrayado la preocupación por la privacidad de los datos estudiantiles y la falta de transparencia en la toma de decisiones automatizadas (UNESCO, 2022; Zawacki-Richter et al., 2019).

En contraparte, la IA ofrece oportunidades para enriquecer el proceso educativo, especialmente en el desarrollo del pensamiento crítico. Desde una perspectiva constructivista, como señalan Luckin et al. (2016), los sistemas de IA que desafían activamente a los estudiantes a evaluar y analizar información diversa no solo fortalecen sus habilidades cognitivas, sino que también promueven la autonomía del aprendizaje. Sin embargo, es necesario abordar estos desarrollos desde un marco ético robusto. Los principios de justicia y autonomía, como los propuestos por Floridi y Cowls (2019), exigen que el diseño de estos sistemas garantice la equidad y el respeto por los derechos de los estudiantes, minimizando los riesgos de sesgos algorítmicos y de vigilancia educativa, problemas que han sido destacados por autores como O'Neil (2016) y Williamson (2019).

Al integrar la IA de manera ética y efectiva en la educación superior, las instituciones no solo abordan los desafíos éticos identificados, sino que también promueven un ambiente donde los estudiantes pueden desarrollar habilidades críticas indispensables para enfrentar los complejos retos del siglo XXI.

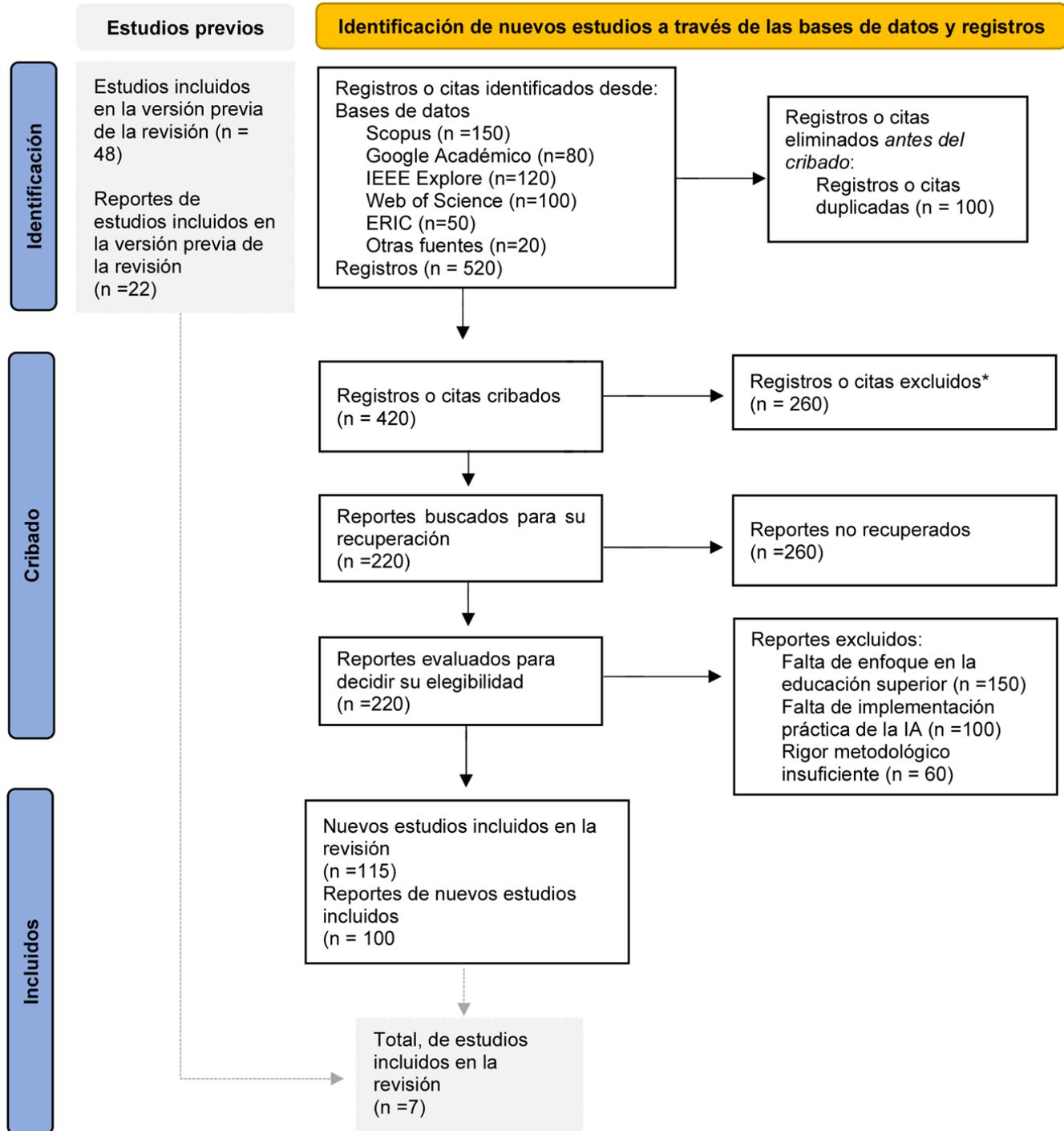
## MÉTODO

Este estudio adopta un enfoque de revisión sistemática de la literatura para analizar el uso de la inteligencia artificial en la educación, con especial atención a las consideraciones éticas y el fomento del pensamiento crítico. Para garantizar una

---

**Figura 1**

Diagrama de flujo PRISMA 2020



Nota: El nuevo diseño se ha adaptado a partir de los diagramas de flujo propuestos por Boers, 2018; Mayo-Wilson et al., 2018, y Stovold et al. (2014)

revisión exhaustiva y rigurosa, se han establecido criterios de inclusión y exclusión que permiten identificar estudios relevantes. Estos criterios incluyeron la selección de investigaciones empíricas y teóricas revisadas por pares, publicadas entre los años 2019 al 2024, que abordaran las implicaciones éticas del uso de IA en educación y su potencial para mejorar el pensamiento crítico. Se excluyeron estudios puramente técnicos que no exploraban estas dimensiones éticas o pedagógicas. El proceso de búsqueda se realizó en bases de datos académicas como ERIC, Google Scholar y JSTOR, empleando términos clave como “inteligencia artificial en educación”, “ética en IA”, “pensamiento crítico” y “tecnologías emergentes en el ámbito educativo”. Por medio del diagrama de flujo PRISMA presentado en la Figura 1 se ilustra el método de revisión sistemática, donde se observa el contenido estructurado del proceso de selección y análisis.

La selección de estudios se realizó mediante la revisión de títulos, resúmenes y, posteriormente, una evaluación detallada de los textos completos, aplicando herramientas de evaluación de calidad, como el Critical Appraisal Skills Programme –CASP–, que asegura que solo los estudios metodológicamente sólidos sean incluidos en el análisis (CASP, 2018). Esta evaluación permitió garantizar la fiabilidad y validez de los estudios seleccionados. Al aplicar la lista de verificación del CASP se puede observar que este estudio cualitativo sobre la IA en educación aborda adecuadamente las consideraciones éticas y ofrece una visión bien fundamentada sobre el impacto de la IA en el pensamiento crítico. Las preguntas del CASP (Tabla 1) nos ayudan a evaluar si los hallazgos del estudio son relevantes, aplicables y están bien sustentados por los datos.

**Tabla 1**  
Análisis CASP

Sección	Pregunta de evaluación
Claridad del propósito	¿Define el propósito en términos de ética y pensamiento crítico en IA educativa?
Relevancia y contexto	¿Contextualiza la necesidad ética en educación? ¿Aborda el impacto en estudiantes y docentes?
Metodología	¿Usa métodos adecuados para evaluar el impacto en el pensamiento crítico?
Resultado y evidencia	¿Presenta datos que respalden los argumentos? ¿Muestra pruebas de desarrollo del pensamiento crítico a través de IA?
Impacto en la comunidad	¿Considera la privacidad y bienestar de estudiantes/docentes? ¿Evalúa la percepción de la comunidad educativa sobre la ética en IA?
Conclusión y recomendaciones	¿Las conclusiones son coherentes con los resultados? ¿Ofrece recomendaciones aplicables para fomentar el pensamiento crítico con IA en educación?

Fuente: Elaboración propia del autor con la adaptación de la metodología CASP.

La revisión también incorporó un análisis comparativo internacional, considerando países como Finlandia, Singapur y Estonia, reconocidos por sus políticas pioneras en la implementación ética de la IA en educación. Este enfoque comparativo resulta

valioso para evaluar cómo las diferentes naciones están enfrentando los desafíos éticos relacionados con la IA, particularmente en términos de equidad y transparencia (OECD, 2021).

Los resultados de la revisión destacan varios desafíos éticos asociados con el uso de la IA en la educación superior. En primer lugar, los sesgos algorítmicos fueron identificados como un problema crítico. Diversos estudios sugieren que los algoritmos, al estar entrenados con grandes cantidades de datos que pueden contener sesgos de género, raza o clase social, tienden a perpetuar y en algunos casos amplificar las desigualdades preexistentes. Esto es especialmente preocupante en contextos educativos, ya que estas tecnologías pueden influir en decisiones importantes, como la admisión de estudiantes o la evaluación académica (Popenici y Kerr, 2017; Williamson, 2019). La investigación refuerza la necesidad de desarrollar algoritmos que sean justos y transparentes, garantizando que las decisiones tomadas por la IA no refuercen las desigualdades estructurales.

Otro de los principales desafíos éticos es la privacidad de los datos de los estudiantes. Las plataformas de IA recopilan grandes cantidades de datos con el objetivo de personalizar el aprendizaje, pero esto plantea serios riesgos para la privacidad de los estudiantes. Williamson (2019) advierte que las plataformas educativas a menudo tratan los datos de los estudiantes como productos comercializables, lo que puede comprometer su autonomía y su derecho a la privacidad. La falta de claridad en cómo se recopilan y utilizan estos datos refuerza la necesidad de políticas sólidas de protección de la privacidad que garanticen la seguridad de la información y el consentimiento informado de los usuarios.

Además, la transparencia en los sistemas de IA es un tema recurrente en la literatura. Los sistemas de IA en la educación suelen ser opacos, lo que significa que las decisiones que toman son difíciles de explicar y auditar. Esta opacidad puede afectar la confianza en la IA y plantear serios problemas de rendición de cuentas. Los estudios revisados coinciden en que es fundamental que las decisiones automatizadas sean comprensibles tanto para los educadores como para los estudiantes, y que existan mecanismos claros para supervisar y corregir posibles errores o sesgos (Decker et al., 2021).

A pesar de estos desafíos, la revisión también pone de relieve el potencial de la IA para personalizar el aprendizaje y mejorar el pensamiento crítico. Holmes, Bialik y Fadel (2021) destacan que la IA permite a los estudiantes avanzar a su propio ritmo, personalizando la experiencia de aprendizaje de acuerdo con sus necesidades individuales, lo que puede fortalecer sus habilidades de análisis y evaluación crítica. Sin embargo, la eficacia de estas tecnologías depende en gran medida de la capacitación docente y de cómo se integran en el currículo educativo. Los docentes necesitan ser formados para utilizar estas herramientas de manera efectiva, ya que la interacción humana sigue siendo clave para el desarrollo del pensamiento crítico (Popenici y Kerr, 2017).

---

Otro aspecto que merece atención es la equidad en el acceso a la IA. A pesar de sus beneficios, no todos los estudiantes tienen acceso equitativo a las tecnologías basadas en IA, especialmente en contextos de bajos recursos. La brecha digital es un desafío ético significativo, ya que la falta de infraestructura tecnológica puede ampliar las desigualdades educativas, lo que subraya la necesidad de políticas que prioricen la equidad en el acceso a la tecnología (OECD, 2021). Países como Finlandia y Singapur han adoptado medidas para garantizar que todos los estudiantes tengan acceso a estas herramientas, sin importar su contexto socioeconómico, lo que puede servir como modelo para otras naciones (OECD, 2021).

## RESULTADOS

Esta investigación no solo identifica los desafíos éticos fundamentales asociados con la IA en la educación superior, sino que también propone estrategias concretas y recomendaciones para mitigar dichos desafíos. Al hacerlo subraya la importancia de un enfoque ético y reflexivo hacia la integración de la IA en el ámbito educativo, promoviendo un entorno educativo más inclusivo y equitativo para todos los estudiantes.

El análisis comparativo del impacto de la IA en la educación debe considerar tanto los beneficios potenciales como la mejora del pensamiento crítico y la personalización del aprendizaje, como los desafíos, incluyendo la equidad en las oportunidades y los riesgos éticos asociados. La combinación de evidencia empírica y estudios de caso permitirá una evaluación completa de cómo las tecnologías de IA están transformando el panorama educativo y cómo se pueden mitigar los riesgos para maximizar sus beneficios.

Durante la revisión sistemática de la literatura se identificaron varios desafíos significativos asociados con el uso de IA en la educación superior. En primer lugar, los estudios han documentado cómo los sesgos algorítmicos presentes en los sistemas de IA pueden intensificar la discriminación en contextos educativos, afectando decisiones clave como la admisión y la evaluación estudiantil (Decker et al., 2021; Popenici y Kerr, 2017). Esto se alinea con las preocupaciones expresadas por Binns (2018), quien argumenta que los sesgos algorítmicos son una consecuencia directa de los datos de entrenamiento y de los diseñadores de los algoritmos, y que estos pueden perpetuar las desigualdades preexistentes en lugar de corregirlas. De manera similar, O'Neil (2016) en su obra *Weapons of math destruction* señala cómo los algoritmos, si no se auditan y supervisan adecuadamente, pueden tener un impacto devastador en la equidad, ya que operan bajo una falsa percepción de objetividad.

Esta preocupación subraya la necesidad urgente de abordar estos sesgos para garantizar una educación más equitativa y justa para todos los estudiantes. Aunque algunos estudios, como el de Holmes et al. (2021), sugieren que los sistemas de IA pueden mejorarse con técnicas de *machine learning* justas, como la incorporación de

criterios de equidad en el diseño de algoritmos, persiste el desafío de equilibrar la innovación tecnológica con la responsabilidad ética.

Además, la protección de la privacidad y seguridad de los datos de los estudiantes ha emergido como otro desafío crucial. La recopilación masiva de datos en plataformas educativas basadas en IA plantea riesgos significativos, y es fundamental establecer prácticas sólidas de manejo de datos que respeten los derechos individuales y la confidencialidad (UNESCO, 2022; Zawacki-Richter et al., 2019). En este contexto, Williamson (2019) sostiene que las plataformas educativas y los sistemas de IA no solo recogen datos sensibles, sino que también pueden convertir a los estudiantes en “productos” cuyo valor reside en los datos que generan. Esto crea un dilema ético, pues mientras la IA promete mejorar el aprendizaje personalizado, también abre la puerta a una vigilancia educativa que podría comprometer la autonomía de los estudiantes.

La falta de transparencia y rendición de cuentas en los algoritmos educativos automatizados también constituye un desafío considerable. La opacidad en los procesos de toma de decisiones puede dificultar la comprensión de cómo se realizan evaluaciones y recomendaciones, lo que podría afectar negativamente la equidad educativa y la confianza pública en estas tecnologías (Decker et al., 2021; Popenici y Kerr, 2017). Al igual que Crawford y Whittaker (2019), quienes abogan por una mayor explicabilidad en los sistemas algorítmicos, el presente estudio refuerza la idea de que los usuarios deben tener acceso a explicaciones claras y comprensibles sobre cómo se generan las decisiones automatizadas, especialmente en el ámbito educativo, donde estas decisiones pueden afectar el futuro académico y profesional de los estudiantes.

Sin embargo, la revisión también reveló oportunidades significativas para un uso ético y efectivo de la IA en la educación superior. Diseñar sistemas de IA que fomenten el pensamiento crítico y la autonomía estudiantil puede contrarrestar los sesgos algorítmicos identificados y mejorar la calidad del aprendizaje (Popenici y Kerr, 2017; Zawacki-Richter et al., 2019). Este enfoque es apoyado por Luckin et al. (2016), quienes subrayan que la IA tiene el potencial de empoderar a los estudiantes, especialmente si se utiliza para desarrollar sus habilidades de análisis crítico y resolución de problemas. A diferencia de los sistemas pasivos, la IA puede ser programada para presentar desafíos cognitivos adaptativos que incentiven el pensamiento crítico y la autonomía.

Implementar políticas claras y códigos de ética robustos es crucial para guiar el desarrollo responsable de la IA, asegurando que estos avances tecnológicos beneficien a todos los estudiantes de manera equitativa (UNESCO, 2022; Zawacki-Richter et al., 2019). Este punto es particularmente relevante si se consideran las recomendaciones de Floridi y Cowls (2019), quienes argumentan que el desarrollo de una IA responsable debe estar basado en principios éticos que vayan más allá de la mera legalidad, abarcando aspectos como la justicia, la autonomía y la privacidad. Esta visión es complementaria a la de Mittelstadt et al. (2016), que propone un marco

de ética algorítmica que permita auditar los sistemas de IA y evaluar su impacto en términos de equidad y transparencia.

La investigación titulada “La inteligencia artificial en educación: consideraciones éticas y fomento del pensamiento crítico” se centra en analizar el uso de la IA en el ámbito educativo, con especial atención a las implicaciones éticas y el potencial de la IA para promover el pensamiento crítico entre los estudiantes. El estudio aborda tres áreas principales: los desafíos éticos de la IA en la educación, las oportunidades que presenta para el fomento del pensamiento crítico y las recomendaciones para un uso responsable de estas tecnologías.

Los resultados muestran que los desafíos éticos más significativos están relacionados con los sesgos algorítmicos, la privacidad y seguridad de los datos, y la transparencia en los sistemas de IA. Estos desafíos tienen un impacto directo en la equidad y justicia del entorno educativo, resaltando la necesidad de políticas y prácticas que mitiguen estos riesgos. Comparativamente, autores como Eubanks (2018) exploran cómo los sistemas algorítmicos en ámbitos como la educación o el empleo pueden perpetuar desigualdades sistémicas si no se controlan adecuadamente. La investigación en cuestión coincide en la urgencia de desarrollar auditorías sistemáticas y prácticas éticas para prevenir estos riesgos y asegurar una implementación más justa y equitativa de la IA en la educación. La literatura sobre la integración de la IA en la educación revela una transformación profunda en el rol del docente, desafiando la noción de que la tecnología lo desplazará. Lejos de ser reemplazado, el docente emerge como un agente aún más crucial en el ecosistema educativo potenciado por la IA, adoptando nuevas funciones y responsabilidades. Como diseñadores curriculares, los docentes aprovechan la capacidad de la IA para personalizar los itinerarios de aprendizaje, seleccionando y organizando recursos digitales y actividades que se alinean con las necesidades individuales de cada estudiante (Holmes et al., 2021). Asimismo, la automatización de tareas repetitivas gracias a la IA libera al docente para que pueda dedicar más tiempo a la interacción individualizada con sus estudiantes, adoptando un rol de mentoría en el que se brinda retroalimentación personalizada y se guía a los estudiantes en el desarrollo de habilidades de orden superior como el pensamiento crítico, la resolución de problemas y la colaboración (Perrenoud, 2010). Sin embargo, esta transformación también exige que el docente asuma la responsabilidad de guardián ético, comprendiendo y abordando los desafíos que plantea la IA en cuanto a sesgos algorítmicos, privacidad de datos y uso responsable de la tecnología (Zawacki-Richter et al., 2019). En definitiva, nuestra revisión sistemática revela que la IA no disminuye la importancia del docente, sino que la amplifica y la redefine. El docente en la era de la IA se convierte en un diseñador, mentor y guardián ético, un agente esencial para asegurar que la tecnología se utilice de manera efectiva y responsable para potenciar el aprendizaje y el desarrollo integral de cada estudiante.

**Tabla 2**

*Principales desafíos éticos en el uso de la IA en educación superior*

Desafío ético	Descripción
Sesgos algorítmicos	Presencia de sesgos de género, raza, clase social, etc., en los datos de entrenamiento y modelos de IA, lo que puede generar discriminación (Decker et al., 2021; Popenici y Kerr, 2017)
Privacidad y seguridad de datos	Preocupaciones sobre la recopilación, almacenamiento y uso de datos personales de estudiantes sin su consentimiento o conocimiento (UNESCO, 2019; Zawacki-Richter et al., 2019)
Transparencia y rendición de cuentas	Falta de explicabilidad y justificación de las decisiones tomadas por los sistemas de IA, dificultando la supervisión y el cuestionamiento (Decker et al., 2021; Popenici y Kerr, 2017)
Comercialización de los datos	Las plataformas educativas que utilizan IA pueden tratar los datos de los estudiantes como un producto comercializable, comprometiendo la privacidad y autonomía de los estudiantes. Esto plantea serias cuestiones éticas sobre los derechos de los estudiantes en cuanto a la propiedad y uso de sus datos (Williamson, 2019)
Rendición de cuentas limitada	La falta de mecanismos robustos para auditar y evaluar los algoritmos de IA en la educación dificulta la supervisión de los impactos de estas tecnologías. Sin auditorías adecuadas, los sesgos algorítmicos pueden pasar desapercibidos, perpetuando desigualdades sin posibilidad de corrección o mejora (Mittelstadt et al., 2016; Eubanks, 2018; Floridi y COWLS, 2019)
Rol del docente	La integración de la inteligencia artificial en la educación no busca reemplazar al docente, sino transformar su rol hacia una posición de mayor importancia como facilitador del aprendizaje y guía ético (Holmes et al., 2021)

*Fuente:* Elaboración propia a partir de datos obtenidos de la revisión de la literatura.

**Tabla 3**

*Estrategias y buenas prácticas para el uso ético de la IA en el fomento del pensamiento crítico*

Estrategias	Descripción
Desarrollo de competencias de pensamiento crítico	Diseño, utilización y evaluación de sistemas de IA que desafíen a los estudiantes a cuestionar, analizar y evaluar información de manera crítica (Popenici y Kerr, 2017; Zawacki-Richter et al., 2019)
Pedagogías centradas en el estudiante	Utilización de la IA para personalizar el aprendizaje y fomentar la autonomía, la autorregulación y la participación activa de los estudiantes (Popenici y Kerr, 2017; UNESCO, 2019)
Colaboración interdisciplinaria	Trabajo conjunto entre expertos en IA, educación y ética para garantizar el desarrollo e implementación responsable de estos sistemas (Decker et al., 2021; UNESCO, 2019)
Evaluación continua	Monitorear y evaluar de manera continua los impactos de la IA en los resultados educativos, así como cualquier efecto adverso (Luckin et al., 2016)

*Fuente:* Elaboración propia a partir de datos obtenidos de la revisión de la literatura.

La Tabla 3 destaca estrategias y buenas prácticas que pueden maximizar los beneficios del uso de la IA en el ámbito educativo, enfocándose especialmente en el desarrollo del pensamiento crítico. Cada estrategia propuesta contribuye de manera significativa a crear un entorno de aprendizaje más eficaz y ético.

Es crucial identificar prácticas efectivas y abordar los riesgos éticos asociados con la IA en la educación. Decker, Peebles y Cole (2021) sugieren que las mejores prácticas incluyen la capacitación continua para los docentes y la implementación de estrategias pedagógicas que complementen las herramientas tecnológicas. La transparencia en el uso de los algoritmos y la auditoría de los sistemas de IA son esenciales para mitigar los sesgos y la discriminación inadvertida (Crawford y Whittaker, 2019). Los estudios recientes y las políticas actualizadas en países como Estonia (Estonian

Government, 2023) proporcionan ejemplos valiosos de cómo abordar estos desafíos y promover una educación más equitativa mediante el uso de IA. Estonia ha actualizado su legislación para mejorar la protección de datos y la seguridad en el uso de IA, reflejando un compromiso continuo con la ética y la innovación en el ámbito educativo (Estonian Data Protection Inspectorate, 2023).

Los hallazgos empíricos respaldan las conclusiones de la revisión de la literatura sobre el potencial de la IA para mejorar el pensamiento crítico; no obstante, también revelan la importancia de superar barreras prácticas para una implementación efectiva. Las instituciones deben adoptar políticas de formación y soporte técnico para educadores, así como fomentar una cultura de apertura hacia la innovación tecnológica, como se muestra en la Tabla 4.

**Tabla 4**

*Recomendaciones de políticas y marcos éticos para el uso de la IA en educación superior*

Recomendaciones	Descripción
Políticas institucionales	Desarrollo e implementación de políticas claras sobre el uso ético de la IA, incluyendo principios de equidad, privacidad y transparencia (UNESCO, 2019; Zawacki-Richter et al., 2019)
Códigos de ética	Establecimiento de códigos de conducta y prácticas éticas que guíen el diseño, implementación y uso de sistemas de IA en el ámbito educativo (UNESCO, 2019; Zawacki-Richter et al., 2019)
Alfabetización en IA	Programas de formación y sensibilización para docentes y estudiantes sobre los usos, beneficios, riesgos y consideraciones éticas de la IA (Popenici y Kerr, 2017; Zawacki-Richter et al., 2019)

*Fuente:* Elaboración propia a partir de datos obtenidos de la revisión de la literatura.

## CONCLUSIONES

La incorporación de la inteligencia artificial –IA– en la educación superior presenta tanto oportunidades transformadoras como desafíos éticos sustanciales. Los resultados de esta revisión sistemática indican que la IA tiene un gran potencial para personalizar el aprendizaje y fomentar el pensamiento crítico, dos habilidades clave en los entornos educativos modernos. Sin embargo, los estudios también destacan los riesgos relacionados con los sesgos algorítmicos, la protección de la privacidad y la transparencia en la toma de decisiones (Floridi y Cowls, 2019; Popenici y Kerr, 2017; Zawacki-Richter et al., 2019). Estos desafíos exigen la implementación de marcos éticos sólidos que guíen el desarrollo y uso de estas tecnologías.

En cuanto a los sesgos algorítmicos, la revisión revela que los algoritmos de IA entrenados con datos sesgados pueden perpetuar o amplificar las desigualdades preexistentes, afectando procesos clave como la admisión y evaluación de los estudiantes (Decker et al., 2021; O’Neil, 2016). Esto subraya la necesidad de desarrollar algoritmos que no solo sean técnicamente avanzados, sino que también garanticen equidad y justicia en su aplicación (Floridi y Cowls, 2019). Además, la protección de los datos sigue siendo una preocupación central, ya que la recopilación masiva de información personal de los estudiantes plantea riesgos significativos para su privacidad

(Williamson, 2019). Los sistemas de IA no deben comprometer la autonomía de los estudiantes ni exponer su información sin un consentimiento adecuado y políticas de protección claras (UNESCO, 2022).

Otro desafío clave identificado es la falta de transparencia en los sistemas de IA, que a menudo funcionan como “cajas negras”, lo que dificulta que los estudiantes y educadores comprendan cómo se toman las decisiones (Decker et al., 2021). Esto afecta no solo la confianza en la IA sino también la capacidad de auditar y corregir posibles errores o sesgos. La literatura insiste en la necesidad de que los sistemas de IA en educación sean explicables y transparentes, de modo que tanto los docentes como los estudiantes puedan comprender y cuestionar las decisiones automatizadas (Crawford y Whittaker, 2019).

A pesar de estos desafíos, los estudios revisados confirman que la IA tiene un potencial considerable para personalizar el aprendizaje y fomentar el pensamiento crítico (Holmes et al., 2021; Luckin et al., 2016). La IA permite a los estudiantes avanzar a su propio ritmo, proporcionando retroalimentación inmediata y adaptativa, lo que puede mejorar su capacidad para analizar, evaluar y reflexionar críticamente sobre la información. No obstante, la efectividad de la IA en este ámbito depende en gran medida de cómo se integra en el currículo y de la capacitación docente para utilizar estas herramientas de manera efectiva (Popenici y Kerr, 2017). Este estudio proporciona un análisis exhaustivo de los desafíos éticos asociados con el uso de la IA en la educación superior, además de destacar su potencial para mejorar el desarrollo del pensamiento crítico. Sin embargo, existen limitaciones. Gran parte de los estudios revisados provienen de contextos educativos de países desarrollados, como Finlandia, Singapur y Estonia (OECD, 2021), lo que restringe la generalización de los resultados a entornos con menos recursos tecnológicos. Además, si bien se ha identificado un impacto positivo de la IA en el pensamiento crítico, faltan estudios empíricos que analicen su efecto a largo plazo en el desarrollo de habilidades cognitivas.

Para futuras investigaciones se sugiere ampliar el análisis del uso de la IA en contextos educativos de países con menos recursos, donde la brecha digital sigue siendo una barrera significativa; además sería esencial realizar estudios longitudinales que evalúen el impacto de la IA en el pensamiento crítico y el desarrollo cognitivo a largo plazo, abordando no solo los efectos inmediatos sino también las implicaciones en la formación integral de los estudiantes (Zawacki-Richter et al., 2019). También es crucial seguir investigando los mecanismos de transparencia y las formas en que los sistemas de IA pueden ser diseñados para ser auditables y comprensibles para todos los usuarios, garantizando la rendición de cuentas en la toma de decisiones automatizadas (Decker et al., 2021). Aunque el estudio tiene algunas limitaciones, es útil para guiar futuras investigaciones y para que los responsables de políticas y diseñadores de herramientas basadas en IA tomen decisiones informadas sobre la integración de estas tecnologías en la educación, para adaptar la estructura al contexto del tipo de investigación.

La IA en la educación superior ofrece grandes oportunidades, pero su implementación debe estar acompañada de un enfoque ético que garantice la equidad, la transparencia y la protección de los derechos de los estudiantes. Las instituciones educativas, junto con los responsables de políticas y los desarrolladores tecnológicos, deben trabajar de manera conjunta para diseñar políticas claras que guíen el uso responsable de la IA en la educación, asegurando que esta tecnología potencie la enseñanza, sin comprometer los principios éticos que rigen la educación.

## REFERENCIAS

- Boers, M. (2018). Graphics and statistics for cardiology: Designing effective tables for presentation and publication. *Heart*, 104(3), 192-200. <https://doi.org/10.1136/heartjnl-2017-311581>
- Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. *Proceedings of the 2018 Conference on Fairness, Accountability, and Transparency* (pp. 149-159).
- CASP [Critical Appraisal Skills Programme] (2018). *CASP Checklists*. <https://casp-uk.net/casp-tools-checklists/>
- Crawford, K., y Whittaker, M. (2019). *AI now 2019 report*. AI Now Institute. <https://ainowinstitute.org/reports.html>
- Decker, A., Peebles, E., y Cole, R. (2021). *AI and ethics in educational contexts: Addressing algorithmic bias and accountability*. AI and Ethics.
- Estonian Data Protection Inspectorate (2023). *Guidelines for data protection in AI applications in education*. <https://www.aki.ee/en>
- Estonian Government (2023). *National strategy for artificial intelligence in education*. Estonian Ministry of Education and Research. <https://www.hm.ee/en>
- Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.
- Floridi, L., y Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1). <https://doi.org/10.1162/99608f92.8cd550d1>
- Holmes, W., Bialik, M., y Fadel, C. (2021). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.
- Luckin, R., Holmes, W., Griffiths, M., y Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson.
- Mayo-Wilson, E., Li, T., Fusco, N., y Dickersin, K. (2018). Practical guidance for using multiple data sources in systematic reviews and meta-analyses (with examples from the MUDS study). *Research Synthesis Methods*, 9(1), 2-12. <https://doi.org/10.1002/jrsm.1277>
- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., y Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2). <https://doi.org/10.1177/2053951716679679>
- OECD [Organization for Economic Co-operation and Development] (2021). *AI principles in education: Addressing equity and access in emerging technologies*. OECD.
- O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown Publishing Group.
- Perrenoud, P. (2010). *Desarrollar la práctica reflexiva en el oficio de enseñar: profesionalización y razón pedagógica*. Graó.

- Popenici, S. A., y Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning*, 12, 22. <https://doi.org/10.1186/s41039-017-0062-8>
- Stovold, E., Beecher, D., Foxlee, R., y Noel-Storr, A. (2014). Study flow diagrams in Cochrane systematic review updates: An adapted PRISMA flow diagram. *Systematic Reviews*, 3, 54. <https://doi.org/10.1186/2046-4053-3-54>
- UNESCO [Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura] (2022). *Recomendación sobre la ética de la inteligencia artificial*. [https://unesdoc.unesco.org/ark:/48223/pf0000381137\\_spa](https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa)
- Williamson, B. (2019). Policy networks, performance metrics and platform markets: Charting the expanding data infrastructure of higher education. *British Journal of Educational Technology*, 50(6), 2794-2809. <https://doi.org/10.1111/bjet.12849>
- Zawacki-Richter, O., Marín, V. I., Bond, M., y 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, 39. <https://doi.org/10.1186/s41239-019-0171-0>

---

Cómo citar este artículo:

Nozato López, M. J. (2024). La inteligencia artificial en educación: consideraciones éticas y fomento al pensamiento crítico. *RECIE. Revista Electrónica Científica de Investigación Educativa*, 8, e2357. <https://doi.org/10.33010/recie.v8i0.2357>



Todos los contenidos de RECIE. *Revista Electrónica Científica de Investigación Educativa* se publican bajo una licencia de Creative Commons Reconocimiento-NoComercial 4.0 Internacional, y pueden ser usados gratuitamente para fines no comerciales, dando los créditos a los autores y a la revista, como lo establece la licencia.

---