MAESTRO y SOCIEDAD e-ISSN 1815-4867

Volumen 23 Número 3 Año 2026
Artículo original

De la respuesta de la IA al juicio propio: pensamiento crítico y autonomía del estudiante

From AI-Generated Responses to Independent Judgment: Critical Thinking and Student Autonomy

Da resposta da IA ao julgamento próprio: pensamento crítico e autonomia do estudante

Juan de Jesús Alvarado Ortiz [1]*, https://orcid.org/ 0000-0002-5304-0757
Oralia Mercedes Moguel Morales [2], https://orcid.org/ 0009-0007-2944-4659
Francisco Jesús Vieyra González [3], https://orcid.org/ 0000-0002-0592-1576
Onelia Hernández Torres [4], https://orcid.org/0009-0004-4757-6362

[1] Universidad Virtual del Estado de Guanajuato (UVEG), México

[2] Escuela Normal Experimental "Fray Matías Antonio de Córdova y Ordóñez", México

[3] Universidad Nacional Autónoma de México (UNAM), México

[4] Escuela Normal de Licenciatura en Educación Primaria "Lic. Manuel Larrainzar", México

*Autor para correspondencia. email jualvaradoo@uveg.edu.mx

Para citar este artículo: Alvarado Ortiz, J. J., Moguel Morales, O. M., Vieyra González, F. J. y Hernández Torres, O. (2026). De la respuesta de la IA al juicio propio: pensamiento crítico y autonomía del estudiante. *Maestro y Sociedad*, 23(3), 3434-3450. https://maestroysociedad.uo.edu.cu

RESUMEN

Introducción: La creciente incorporación de la inteligencia artificial generativa en la educación superior ha transformado la manera en que los estudiantes acceden, procesan y producen información, generando nuevas oportunidades de aprendizaje, pero también riesgos asociados con la dependencia cognitiva. El objetivo del estudio fue analizar la relación entre el uso de inteligencia artificial generativa, el pensamiento crítico y la autonomía del estudiante, identificando las condiciones que favorecen el juicio propio y aquellas que pueden promover la delegación excesiva del razonamiento. Materiales y Métodos: Se desarrolló una revisión sistemática de literatura bajo las directrices PRISMA 2020, considerando publicaciones entre 2021 y 2026 en bases de datos académicas de alto impacto. Tras el proceso de identificación, cribado y elegibilidad, se incluyeron 24 estudios en la síntesis cualitativa. Los resultados evidenciaron cuatro ejes principales: apoyo cognitivo de la IA, pensamiento crítico y evaluación de respuestas, autorregulación y agencia estudiantil, y dependencia o descarga cognitiva. Conclusiones: Se concluye que la IA no fortalece ni debilita automáticamente la autonomía; su efecto depende del control cognitivo que conserve el estudiante. Cuando las respuestas son contrastadas, verificadas y reflexionadas, la IA puede potenciar el aprendizaje; cuando se aceptan acríticamente, aumenta el riesgo de dependencia y pérdida de juicio propio.

Palabras clave: Inteligencia artificial generativa, pensamiento crítico, autonomía estudiantil, autorregulación del aprendizaje, dependencia cognitiva.

ABSTRACT

Introduction: The growing integration of generative artificial intelligence into higher education has transformed the way students access, process, and produce information, creating new learning opportunities while also raising risks associated with cognitive dependence. The objective of this study was to analyze the relationship between the use of generative artificial intelligence, critical thinking, and student autonomy, identifying the conditions that foster independent judgment and those that may promote excessive delegation of reasoning. Materials and methods: A systematic literature review was conducted following the PRISMA 2020 guidelines, considering publications from 2021

Licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Internacional

to 2026 in high-impact academic databases. After the identification, screening, and eligibility stages, 24 studies were included in the qualitative synthesis. The results revealed four main thematic areas: AI as cognitive support, critical thinking and evaluation of AI-generated responses, self-regulation and student agency, and cognitive dependence or cognitive offloading. Conclusions: It is concluded that AI does not automatically strengthen or weaken student autonomy; its effect depends on the degree of cognitive control retained by the learner. When AI-generated responses are contrasted, verified, and critically reflected upon, AI can enhance learning; when they are accepted uncritically, the risk of dependence and loss of independent judgment increases.
Keywords: Generative artificial intelligence, critical thinking, student autonomy, self-regulated learning, cognitive dependence.

RESUMO

Introdução: La crescente incorporação da inteligência artificial generativa no ensino superior tem transformado a maneira como os estudantes acessam, processam e produzem informações, criando novas oportunidades de aprendizagem, mas também riscos associados à dependência cognitiva. O objetivo deste estudo foi analisar a relação entre o uso da inteligência artificial generativa, o pensamento crítico e a autonomia do estudante, identificando as condições que favorecem o julgamento próprio e aquelas que podem promover a delegação excessiva do raciocínio. Foi realizada uma revisão sistemática da literatura seguindo as diretrizes PRISMA 2020, considerando publicações entre 2021 e 2026 em bases de dados acadêmicas de alto impacto. Após as etapas de identificação, triagem e elegibilidade, 24 estudos foram incluídos na síntese qualitativa. Os resultados evidenciaram quatro eixos principais: IA como apoio cognitivo, pensamento crítico e avaliação de respostas, autorregulação e agência estudantil, e dependência ou descarga cognitiva. Conclui-se que a IA não fortalece nem enfraquece automaticamente a autonomia; seu efeito depende do grau de controle cognitivo mantido pelo estudante. Quando as respostas são contrastadas, verificadas e analisadas criticamente, a IA pode potencializar a aprendizagem; quando aceitas de forma acrítica, aumenta o risco de dependência e perda do julgamento próprio.
Palavras-chave: inteligência artificial generativa, pensamento crítico, autonomia estudantil, aprendizagem autorregulada, dependência cognitiva.

Recibido: 30/3/2026 Aprobado: 15/8/2026

Introducción

La irrupción de la inteligencia artificial generativa (IAG) en los entornos educativos ha modificado de manera significativa la relación de los estudiantes con la información, el conocimiento y la resolución de problemas. Herramientas basadas en grandes modelos de lenguaje, como ChatGPT y otros asistentes conversacionales, permiten obtener explicaciones, resumir contenidos, generar ideas, elaborar textos y recibir retroalimentación prácticamente inmediata (Santana González, 2025). Esta capacidad ha abierto nuevas posibilidades para personalizar el aprendizaje, ampliar el acceso al conocimiento y acompañar procesos académicos que anteriormente dependían casi exclusivamente de la interacción entre docentes, estudiantes y recursos bibliográficos (Sardiñas Padilla & Valdés García, 2025). Sin embargo, la facilidad con la que estas tecnologías producen respuestas coherentes y aparentemente fundamentadas también introduce una cuestión educativa de mayor profundidad: qué ocurre con el razonamiento del estudiante cuando una parte creciente del proceso cognitivo puede ser delegada a una herramienta algorítmica (Padilla Hernández, 2026).

La percepción estudiantil frente a estas tecnologías suele ser favorable, especialmente por su accesibilidad, rapidez y capacidad de brindar apoyo personalizado. Chan y Hu (2023), al analizar las experiencias de estudiantes universitarios, identificaron una actitud generalmente positiva hacia la incorporación de la IA generativa en los procesos de enseñanza y aprendizaje, aunque acompañada por preocupaciones relacionadas con su precisión, privacidad, ética y posibles efectos sobre el desarrollo personal. Este escenario demuestra que la discusión ya no puede limitarse a determinar si la IA debe o no formar parte de la educación superior (Zapata Muriel et al., 2024). Su utilización se encuentra progresivamente incorporada en las prácticas académicas y, por tanto, el desafío se desplaza hacia la manera en que los estudiantes interactúan intelectualmente con sus respuestas y hacia el tipo de aprendizaje que se construye a partir de esa interacción (Trombetta Silvero et al., 2026).

En este contexto, el pensamiento crítico adquiere una relevancia renovada. La capacidad de analizar argumentos, valorar evidencias, reconocer inconsistencias, identificar sesgos y emitir juicios fundamentados

resulta especialmente necesaria cuando los estudiantes trabajan con sistemas capaces de producir información lingüísticamente convincente, pero no necesariamente verdadera, suficiente o contextualizada. Larson et al. (2024) sostienen que las oportunidades y los riesgos derivados de la IA generativa hacen necesario prestar una atención creciente al pensamiento crítico dentro de los procesos educativos. Los autores lo vinculan tanto con la capacidad individual de recurrir a la lógica, la evidencia y el análisis para sustentar afirmaciones como con la disposición para cuestionar perspectivas establecidas (Padilla Hernández, 2024). En este sentido, la alfabetización en IA no debería reducirse al dominio instrumental de una herramienta o a la formulación eficiente de instrucciones, sino incluir la capacidad del estudiante para interrogar los resultados obtenidos, determinar su confiabilidad y decidir cuándo aceptarlos, modificarlos o rechazarlos.

La evidencia disponible muestra, además, que la influencia de la IA sobre las capacidades cognitivas no puede interpretarse de manera lineal. Essel et al. (2024), mediante un estudio desarrollado con estudiantes universitarios, analizaron los efectos de ChatGPT sobre el pensamiento crítico, creativo y reflexivo, encontrando que su incorporación puede producir cambios apreciables en estas habilidades. Estos resultados sugieren que las herramientas generativas pueden convertirse en estímulos para procesos cognitivos complejos cuando son utilizadas dentro de experiencias de aprendizaje estructuradas. Desde esta perspectiva, solicitar explicaciones alternativas, comparar respuestas, identificar errores, construir contraargumentos o justificar por qué determinada solución resulta más adecuada puede transformar la interacción con la IA en una oportunidad para profundizar el razonamiento (Camastra & González Vallejo, 2025). La tecnología, por tanto, no necesariamente desplaza el pensamiento: bajo determinadas condiciones pedagógicas puede convertirse en un interlocutor que lo provoque.

No obstante, esta posibilidad convive con un riesgo igualmente importante: que la asistencia tecnológica sustituya progresivamente actividades cognitivas que el estudiante debería aprender a realizar por sí mismo. La rapidez con la que una herramienta generativa produce una explicación o una solución puede favorecer la aceptación inmediata de su contenido, reduciendo el tiempo dedicado a buscar evidencias, contrastar perspectivas, formular hipótesis y elaborar conclusiones propias. Esta tensión adquiere especial relevancia a partir del trabajo de Darvishi et al. (2024), quienes estudiaron el impacto de la asistencia basada en IA sobre la agencia estudiantil. Sus resultados mostraron una tendencia de los participantes a "rely on rather than learn from AI assistance" [depender de la asistencia de la IA en lugar de aprender de ella] (Darvishi et al., 2024). Este hallazgo introduce una distinción fundamental: obtener mejores resultados mientras la herramienta está disponible no significa necesariamente haber desarrollado las capacidades necesarias para actuar con autonomía cuando dicha asistencia desaparece.

La autonomía del estudiante debe entenderse, en consecuencia, más allá de la posibilidad de aprender sin la presencia constante del docente. En ecosistemas educativos mediados por inteligencia artificial, ser autónomo implica también conservar el control intelectual sobre las decisiones que intervienen en el aprendizaje: formular preguntas relevantes, seleccionar fuentes, determinar qué información merece confianza, supervisar la propia comprensión, identificar errores, revisar argumentos y asumir la responsabilidad final sobre las conclusiones alcanzadas. La agencia estudiantil se relaciona precisamente con esa capacidad de actuar intencionalmente dentro del proceso educativo. Darvishi et al. (2024) advierten que, aunque las tecnologías basadas en IA pueden automatizar y apoyar actividades de aprendizaje, sus efectos sobre la agencia y la capacidad de autorregulación todavía requieren mayor comprensión. Como señalan los propios autores, "students tended to rely on rather than learn from AI assistance", resultado que obliga a reconsiderar modelos educativos donde la eficiencia proporcionada por el algoritmo pueda confundirse con aprendizaje genuino.

El problema resulta especialmente significativo porque una respuesta generada por IA puede ser correcta, parcialmente correcta, sesgada o incluso completamente errónea y, aun así, presentarse mediante una estructura discursiva convincente. Frente a esta condición, el valor educativo ya no reside únicamente en que el estudiante sea capaz de obtener una respuesta, sino en que pueda juzgarla. Esto supone pasar de una lógica centrada en preguntar a la IA "¿cuál es la respuesta?" hacia otra orientada a cuestiones como "¿por qué debería aceptar esta respuesta?", "¿qué evidencias la respaldan?", "¿qué perspectivas han sido omitidas?", "¿podría existir otra explicación?" o "¿puedo defender esta conclusión sin depender de la herramienta?". Desde esta perspectiva, el pensamiento crítico funciona como un mecanismo de mediación entre la información producida algorítmicamente y la construcción de conocimiento por parte del estudiante.

Esta distinción permite reconocer dos posibles trayectorias educativas. En una primera, la respuesta proporcionada por la IA es aceptada con escaso cuestionamiento y utilizada como sustituto del esfuerzo

cognitivo; la tecnología asume progresivamente funciones de búsqueda, interpretación, argumentación y decisión que antes correspondían al estudiante. En una segunda trayectoria, la respuesta generada constituye únicamente un punto de partida que debe ser contrastado, interrogado, reconstruido y sometido a evaluación. En este último escenario, la interacción humano–IA puede favorecer procesos de reflexión, autorregulación y aprendizaje profundo. En consecuencia, la cuestión central no es determinar si la inteligencia artificial piensa por el estudiante, sino establecer qué decisiones cognitivas continúa ejerciendo el estudiante cuando piensa con inteligencia artificial.

A pesar del creciente volumen de investigaciones sobre ChatGPT y otras herramientas generativas en educación superior, todavía existe una fragmentación entre los estudios centrados en aceptación tecnológica, desempeño académico, integridad, pensamiento crítico, autorregulación y agencia estudiantil. La relación entre estos últimos constructos requiere especial atención, puesto que el incremento de la asistencia automatizada puede generar simultáneamente oportunidades para ampliar las capacidades intelectuales y condiciones para incrementar la dependencia cognitiva. Larson et al. (2024) destacan precisamente la necesidad de desarrollar una agenda de investigación más rigurosa sobre las interrelaciones entre el uso de IA generativa y el pensamiento crítico, así como sobre las estrategias educativas capaces de reducir sus amenazas y aprovechar sus beneficios.

El problema que orienta este estudio surge, por tanto, de una paradoja educativa contemporánea: mientras la inteligencia artificial incrementa la capacidad del estudiante para obtener respuestas, no garantiza que aumente proporcionalmente su capacidad para comprenderlas, evaluarlas y tomar decisiones intelectuales independientes. Cuando la producción de contenidos, argumentos y soluciones puede delegarse con facilidad, existe el riesgo de confundir disponibilidad de información con conocimiento, fluidez discursiva con comprensión y eficiencia con autonomía. El desafío para la educación superior consiste en evitar que la IA se convierta en un sustituto silencioso del razonamiento y transformarla, en cambio, en un recurso que obligue al estudiante a argumentar mejor, verificar con mayor rigor y asumir una posición consciente frente a la información que recibe.

En este marco, el objetivo del presente estudio es analizar críticamente la relación entre el uso de inteligencia artificial generativa, el desarrollo del pensamiento crítico y la autonomía del estudiante, identificando las condiciones bajo las cuales la interacción con sistemas de IA puede favorecer la evaluación reflexiva, la autorregulación y la construcción de juicio propio, así como los riesgos asociados con la dependencia y la delegación excesiva de procesos cognitivos. Desde esta perspectiva, se busca contribuir a una comprensión pedagógica de la inteligencia artificial que trascienda la obtención inmediata de respuestas y sitúe nuevamente al estudiante como agente responsable de interpretar, cuestionar, decidir y construir conocimiento.

MATERIALES Y MÉTODOS

Diseño y enfoque del estudio

La investigación se desarrolló mediante una revisión sistemática de la literatura con enfoque cualitativo e interpretativo, orientada a analizar la relación entre el uso de inteligencia artificial generativa, el pensamiento crítico y la autonomía del estudiante en contextos educativos. Para garantizar transparencia, trazabilidad y rigor en el proceso de identificación y selección de los estudios, se siguieron las recomendaciones establecidas por la declaración PRISMA 2020.

La elección de este diseño respondió a la necesidad de integrar evidencia reciente procedente de diferentes investigaciones y, al mismo tiempo, superar una aproximación meramente descriptiva del fenómeno. En este sentido, la revisión buscó comprender no solo los beneficios asociados al empleo de herramientas de inteligencia artificial, sino también las condiciones bajo las cuales estas tecnologías pueden favorecer procesos de reflexión, autorregulación y juicio propio, o, por el contrario, promover dependencia y delegación excesiva de funciones cognitivas.

La unidad de análisis estuvo constituida por investigaciones científicas relacionadas con inteligencia artificial generativa aplicada al ámbito educativo, particularmente aquellas que abordaron dimensiones vinculadas con pensamiento crítico, autonomía, agencia estudiantil, metacognición, autorregulación del aprendizaje y dependencia cognitiva.

Fuentes de información y estrategia de búsqueda

La búsqueda bibliográfica se realizó en Scopus, Web of Science, ERIC y ScienceDirect, complementada con SpringerLink, Wiley Online Library y Taylor & Francis. Se emplearon términos en inglés y español relacionados con inteligencia artificial generativa, pensamiento crítico, autonomía, agencia estudiantil, autorregulación, metacognición y descarga cognitiva. Las palabras clave se combinaron mediante los operadores booleanos AND y OR para ampliar y delimitar la recuperación de estudios pertinentes al objetivo de la revisión.

Se priorizaron publicaciones comprendidas entre 2021 y 2026, prestando especial atención a los estudios publicados desde 2023, debido al crecimiento acelerado de la investigación educativa asociada a herramientas de inteligencia artificial generativa.

Criterios de elegibilidad

Para garantizar la pertinencia de la evidencia seleccionada se establecieron criterios previos de inclusión y exclusión.

- Se incluyeron artículos científicos revisados por pares, publicados en revistas académicas indexadas, escritos en inglés o español y vinculados directamente con el uso de inteligencia artificial en procesos de enseñanza y aprendizaje. Asimismo, se consideraron estudios que analizaran al menos una de las dimensiones centrales de la investigación: pensamiento crítico, autonomía, agencia estudiantil, autorregulación, metacognición, aprendizaje profundo o dependencia cognitiva.
- Se excluyeron documentos duplicados, publicaciones sin revisión por pares, editoriales, opiniones, notas periodísticas, documentos exclusivamente técnicos sobre desarrollo de algoritmos y estudios cuyo foco no guardara relación directa con procesos educativos. También se descartaron investigaciones que abordaran inteligencia artificial únicamente desde perspectivas administrativas, institucionales o tecnológicas, sin analizar su interacción con el aprendizaje del estudiante.

Proceso de selección de los estudios

La selección de los documentos se desarrolló progresivamente siguiendo las fases establecidas por PRISMA 2020: identificación, cribado, evaluación de elegibilidad e inclusión.

En una primera fase se recuperaron los registros mediante las estrategias de búsqueda definidas en las bases de datos seleccionadas. Posteriormente se procedió a identificar y eliminar los documentos duplicados. Los registros restantes fueron sometidos a una primera revisión basada en títulos y resúmenes, mediante la cual se descartaron aquellos que no mantenían una relación directa con el propósito del estudio.

En una segunda fase, los documentos potencialmente pertinentes fueron revisados a texto completo. Durante este proceso se evaluó su correspondencia con los criterios de elegibilidad previamente establecidos, prestando especial atención a la población analizada, el contexto educativo, el tipo de inteligencia artificial utilizada y las variables cognitivas o autorregulatorias examinadas.

Finalmente, los estudios que cumplieron con los criterios de inclusión conformaron el corpus documental utilizado para la síntesis cualitativa. El proceso completo de identificación y selección será representado posteriormente mediante un diagrama de flujo PRISMA, en el que se especificarán los registros encontrados, eliminados, evaluados y finalmente incluidos en la revisión. El proceso de identificación y selección de los estudios se desarrolló de manera secuencial conforme a las fases propuestas por PRISMA 2020. La búsqueda inicial permitió identificar 378 registros procedentes de las diferentes fuentes de información consultadas. Tras la eliminación de 86 documentos duplicados, permanecieron 292 registros, los cuales fueron examinados mediante la revisión de títulos y resúmenes. En esta etapa se excluyeron 218 publicaciones por no guardar una relación suficiente con el objetivo de la investigación.

Posteriormente, se buscaron 74 documentos para su recuperación a texto completo; tres de ellos no pudieron ser obtenidos, por lo que 71 estudios fueron sometidos a evaluación de elegibilidad. Después de aplicar los criterios establecidos, se excluyeron 47 documentos debido principalmente a su relación indirecta con el tema, ausencia de análisis sobre pensamiento crítico o autonomía, falta de correspondencia con el ámbito de la educación superior, carácter no empírico o presencia de información redundante. Como resultado, 24 estudios conformaron el corpus final de la revisión sistemática y fueron incluidos en la síntesis cualitativa. La Figura 1 presenta de manera gráfica el proceso seguido para la selección de la evidencia.

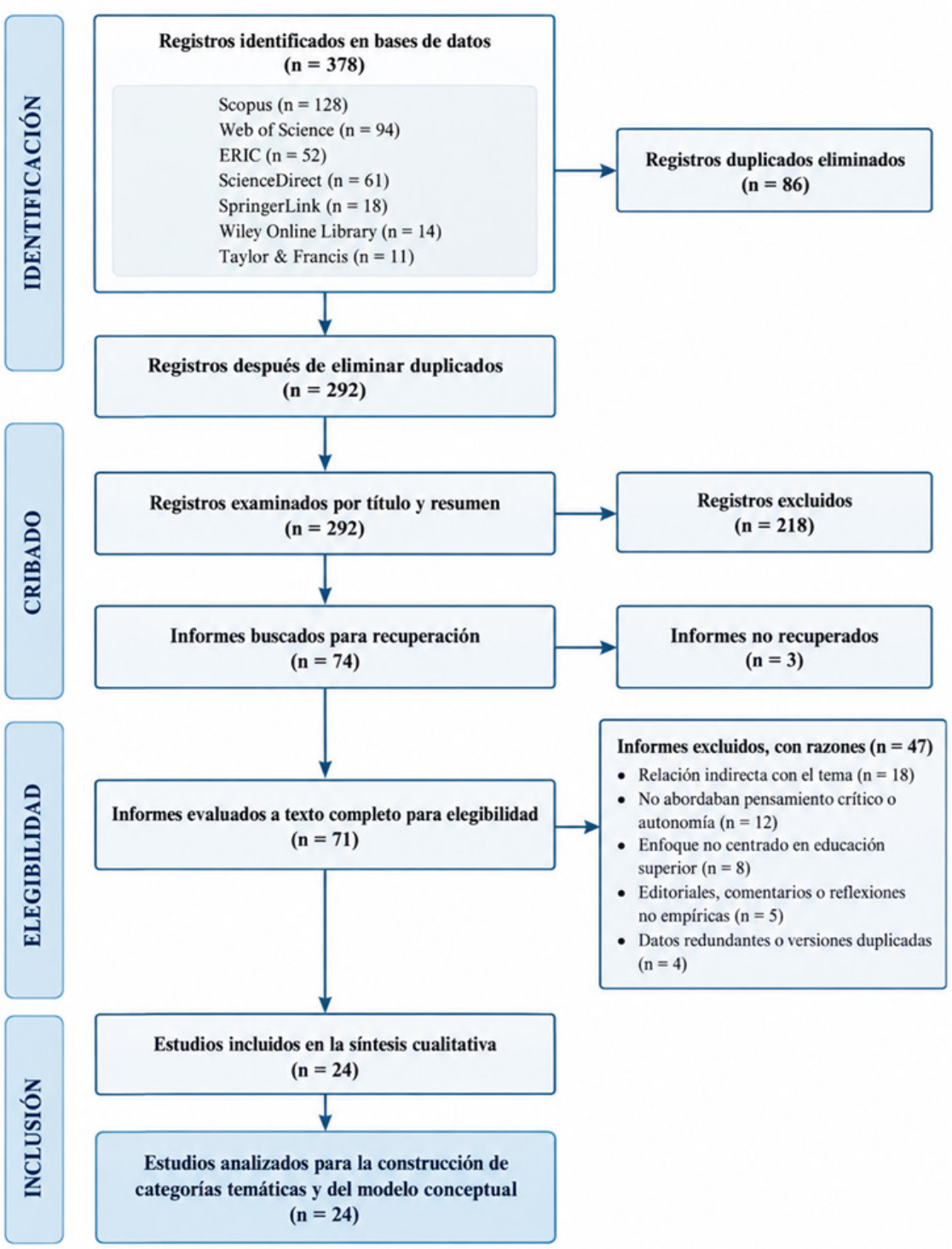


Figura 1. Diagrama de flujo PRISMA 2020 del proceso de selección de estudios.

Fuente: Elaboración propia.

Como se observa en la Figura 1, el proceso de depuración permitió reducir progresivamente el conjunto inicial de documentos hasta obtener una selección final estrechamente vinculada con el propósito de la investigación. La exclusión de estudios no se realizó únicamente en función de criterios formales, sino también considerando su capacidad para aportar evidencia sobre la interacción entre inteligencia artificial generativa y procesos cognitivos vinculados con el pensamiento crítico, la autorregulación, la agencia y la autonomía del estudiante.

Los 24 estudios finalmente incluidos constituyeron la base para la síntesis temática de la evidencia. A partir de ellos se analizaron tanto las posibilidades de la inteligencia artificial como herramienta de apoyo al razonamiento y al aprendizaje autónomo, como los riesgos derivados de la aceptación acrítica de sus respuestas, la delegación del esfuerzo cognitivo y la posible generación de dependencia tecnológica. Esta

selección permitió disponer de un corpus suficientemente específico para examinar las condiciones que pueden conducir desde la utilización de una respuesta generada por IA hacia la construcción de un juicio propio por parte del estudiante.

Extracción y organización de la información

Para sistematizar la evidencia se diseñó una matriz de extracción de datos en la que se registraron las principales características de cada estudio. La matriz consideró variables como autor y año de publicación, país o contexto de estudio, diseño metodológico, población o muestra, herramienta de inteligencia artificial utilizada, dimensión analizada, principales resultados y aportes relacionados con pensamiento crítico y autonomía.

Además, se registraron aquellos hallazgos relacionados con efectos positivos, riesgos o condiciones mediadoras de la interacción entre estudiantes e inteligencia artificial. Este procedimiento permitió comparar resultados provenientes de metodologías y contextos diferentes, identificar coincidencias y reconocer posibles contradicciones dentro de la literatura revisada. La organización de los estudios no se limitó a contabilizar la frecuencia de determinados hallazgos, sino que buscó comprender el significado pedagógico de las relaciones identificadas entre IA, cognición y autonomía.

Análisis y síntesis de la evidencia

La información recopilada fue examinada mediante una síntesis temática cualitativa, orientada a identificar patrones recurrentes, relaciones conceptuales y tendencias emergentes entre los estudios seleccionados. A partir de una lectura comparativa de los documentos, los hallazgos se organizaron inicialmente en cuatro ejes analíticos: IA como apoyo cognitivo para el aprendizaje; pensamiento crítico y evaluación de respuestas generadas por IA; autonomía, autorregulación y agencia estudiantil; y dependencia, descarga cognitiva y pérdida potencial de control intelectual.

Estos ejes permitieron analizar la inteligencia artificial no como una variable aislada, sino como un elemento que interactúa con las decisiones, estrategias y capacidades del estudiante. Particularmente, se examinó si las herramientas generativas promovían procesos de contraste, verificación, reflexión y construcción argumentativa o si, por el contrario, favorecían la aceptación automática de respuestas y la sustitución progresiva del esfuerzo cognitivo.

La comparación entre estudios permitió establecer relaciones entre estos comportamientos y construir una interpretación integradora del fenómeno. De esta manera, la revisión buscó avanzar desde la descripción de las herramientas hacia la comprensión de los procesos cognitivos que determinan si la inteligencia artificial actúa como soporte del aprendizaje o como sustituto parcial del razonamiento.

Consideraciones de rigor metodológico

Para fortalecer la consistencia de la revisión, los criterios de búsqueda, selección y análisis fueron establecidos previamente y aplicados de manera uniforme durante el proceso de revisión. Se procuró mantener correspondencia entre el objetivo del estudio, las fuentes consultadas, los términos de búsqueda, los criterios de elegibilidad y las categorías utilizadas para la síntesis.

Asimismo, se priorizó literatura reciente, revisada por pares y procedente de revistas científicas de reconocida visibilidad académica. La utilización de PRISMA 2020 permitió estructurar de manera transparente el proceso de recuperación y depuración de la evidencia, mientras que la síntesis temática facilitó una interpretación más profunda de los resultados. Este procedimiento metodológico permitió abordar el fenómeno desde una perspectiva equilibrada, reconociendo tanto las posibilidades educativas de la inteligencia artificial como los riesgos asociados con su uso acrítico. Así, el análisis se orientó a comprender una cuestión central: en qué condiciones la interacción con la inteligencia artificial contribuye al desarrollo del juicio propio del estudiante y bajo cuáles puede incrementar su dependencia cognitiva.

RESULTADOS

La revisión sistemática permitió identificar un conjunto de hallazgos convergentes respecto a la relación entre inteligencia artificial generativa, pensamiento crítico y autonomía del estudiante. Los 24 estudios incluidos en la síntesis cualitativa mostraron que el impacto educativo de la IA no depende únicamente de la disponibilidad de la tecnología, sino principalmente de la forma en que el estudiante interactúa con ella, del nivel de acompañamiento pedagógico y de las estrategias metacognitivas que regulan su utilización.

El análisis temático permitió organizar la evidencia en cuatro grandes categorías: a) inteligencia artificial como apoyo cognitivo para el aprendizaje; b) pensamiento crítico y evaluación de las respuestas generadas por IA; c) autonomía, autorregulación y agencia estudiantil; y d) dependencia, descarga cognitiva y pérdida potencial de control intelectual. Estas categorías no se comportaron de manera aislada, sino que mostraron relaciones constantes entre sí, permitiendo reconocer dos formas diferenciadas de interacción con la IA: una orientada al fortalecimiento del aprendizaje y otra caracterizada por una creciente delegación del esfuerzo cognitivo.

La caracterización de los estudios incluidos permitió identificar la diversidad de enfoques metodológicos, contextos educativos y dimensiones analizadas en torno al uso de inteligencia artificial generativa. Con el propósito de organizar la evidencia de manera sistemática, se consideraron aspectos como autor y año de publicación, país o contexto de desarrollo, diseño metodológico, población o muestra, nivel educativo, herramienta de inteligencia artificial empleada, variable principal y hallazgos relevantes. Esta organización facilitó la comparación entre investigaciones y permitió reconocer los principales patrones presentes en la literatura seleccionada. La Tabla 1 sintetiza las características generales de los 24 estudios incluidos en la revisión sistemática.

Tabla 1 Características generales de los 24 estudios incluidos en la revisión sistemática.

Autor/año	Diseño	Contexto educativo	IA utilizada	Variable principal	Hallazgo principal
Chan & Hu (2023)	Estudio empírico	Educación superior	IA generativa/ ChatGPT	Percepciones y uso	Los estudiantes reconocen beneficios de accesibilidad y apoyo, pero expresan preocupaciones sobre precisión, ética y confiabilidad.
Darvishi et al. (2024)	Experimental	Educación superior	Asistencia basada en IA	Agencia estudiantil	La asistencia mejora el desempeño inmediato, pero puede generar dependencia cuando el estudiante se apoya en la IA en lugar de aprender.
Ng et al. (2024)	Estudio piloto	Educación superior	ChatGPT	Autorregulación	El uso guiado de ChatGPT puede favorecer procesos de aprendizaje autorregulado.
Xu et al. (2025)	Cuasiexperimental	Educación superior	IA generativa	Metacognición y autorregulación	El apoyo metacognitivo mejora la capacidad del estudiante para regular su aprendizaje en entornos con IA.
Urban et al. (2025)	Estudio empírico	Escritura académica	ChatGPT	Precisión metacognitiva	Los estudiantes no siempre identifican adecuadamente errores generados por IA, incluso cuando saben que pueden existir.

Fuente. Elaboración de los autores a partir de los estudios incluidos en la revisión sistemática.

Como se observa en la Tabla 1, los estudios incluidos presentan una marcada heterogeneidad metodológica y temática, lo que refleja la amplitud con la que actualmente se investiga la relación entre inteligencia artificial generativa y aprendizaje. Las investigaciones abordan dimensiones como pensamiento crítico, autorregulación, metacognición, agencia estudiantil, autonomía y descarga cognitiva, mediante diseños experimentales, cuasiexperimentales, cualitativos, revisiones y otros enfoques de análisis.

Esta diversidad permitió examinar el fenómeno desde perspectivas complementarias y reconocer que los efectos de la IA no dependen únicamente de la herramienta utilizada, sino también del contexto educativo, del tipo de actividad propuesta y del nivel de participación cognitiva que conserva el estudiante. A partir de esta caracterización, los hallazgos fueron organizados posteriormente en categorías temáticas que permiten comprender con mayor profundidad las condiciones bajo las cuales la IA puede actuar como apoyo al aprendizaje o, por el contrario, favorecer procesos de dependencia y delegación cognitiva.

Inteligencia artificial como apoyo cognitivo para el aprendizaje

Una primera tendencia observada en los estudios analizados fue el reconocimiento de la IA generativa como una herramienta capaz de ofrecer apoyo inmediato durante distintas fases del aprendizaje. Entre sus aportes se identificaron la generación de explicaciones, reformulación de contenidos, retroalimentación, producción de ejemplos, organización de ideas y acompañamiento en la resolución de problemas.

En este sentido, la evidencia muestra que la inteligencia artificial puede funcionar como un recurso de andamiaje cuando el estudiante mantiene una participación en el proceso. Su utilidad resulta especialmente visible cuando las respuestas generadas se emplean como punto de partida para comparar alternativas, reformular preguntas, profundizar conceptos o recibir retroalimentación adicional.

Los estudios también evidenciaron que la interacción con IA puede favorecer experiencias de aprendizaje más personalizadas, debido a la posibilidad de ajustar las consultas al ritmo, nivel de conocimiento y necesidades particulares del estudiante. Sin embargo, este beneficio no se observó de forma automática. Los resultados sugieren que el aprendizaje mejora cuando la asistencia tecnológica se integra dentro de actividades que exigen análisis, revisión y toma de decisiones, y no cuando la herramienta se utiliza exclusivamente para producir una respuesta final.

En consecuencia, la IA mostró mayor potencial educativo cuando cumplió una función de apoyo al razonamiento, en lugar de asumir por completo la resolución de la tarea.

Pensamiento crítico y evaluación de las respuestas generadas por IA

El segundo eje identificado correspondió al pensamiento crítico. Los estudios revisados coincidieron en señalar que uno de los principales desafíos de la IA generativa reside en la capacidad de producir respuestas lingüísticamente coherentes y convincentes, aun cuando estas puedan contener información incompleta, sesgada o incorrecta.

Ante esta situación, la evidencia destacó la necesidad de que el estudiante desarrolle capacidades para contrastar información, verificar fuentes, identificar inconsistencias, evaluar argumentos y determinar la pertinencia de las respuestas obtenidas. En este sentido, el pensamiento crítico aparece como una competencia mediadora entre la producción automatizada de información y la construcción de conocimiento.

Los resultados también permitieron observar que las actividades que exigen al estudiante comparar una respuesta generada por IA con otras fuentes, justificar su aceptación o rechazo, detectar errores o construir contraargumentos producen una interacción cognitivamente más exigente que aquellas centradas exclusivamente en la generación automática de contenidos.

Por tanto, el potencial educativo de la IA parece aumentar cuando se transforma la lógica de uso desde “obtener una respuesta” hacia “evaluar una respuesta”. Este desplazamiento resulta especialmente relevante porque obliga al estudiante a recuperar conocimientos previos, establecer criterios de valoración y asumir una posición frente a la información generada.

La revisión mostró, además, que la alfabetización en IA y el pensamiento crítico se encuentran estrechamente relacionados. El dominio instrumental de una herramienta no garantiza por sí mismo una utilización académicamente adecuada. Los estudiantes requieren comprender las limitaciones de estos sistemas y reconocer que una respuesta plausible no constituye necesariamente una respuesta válida.

Autonomía, autorregulación y agencia estudiantil

La tercera categoría agrupó los estudios relacionados con autonomía, autorregulación y agencia. Los hallazgos evidenciaron que la inteligencia artificial puede contribuir al aprendizaje autónomo cuando permite que el estudiante planifique tareas, solicite retroalimentación, identifique dificultades y ajuste sus estrategias de aprendizaje.

No obstante, la presencia de IA no produjo necesariamente mayores niveles de autonomía. La evidencia mostró que la autorregulación depende en gran medida de la capacidad del estudiante para establecer objetivos, monitorear su comprensión y evaluar críticamente tanto su propio desempeño como la asistencia recibida.

En esta línea, Xu et al. (2025) analizaron específicamente el papel del apoyo metacognitivo en entornos de IA generativa. Su estudio cuasiexperimental con estudiantes universitarios mostró que la incorporación explícita de estrategias metacognitivas constituye un elemento relevante para fortalecer el aprendizaje autorregulado en este tipo de entornos. Los autores destacan que los estudiantes pueden encontrar dificultades para mantener procesos de autorregulación cuando utilizan IA sin orientación específica, mientras que los apoyos metacognitivos favorecen una interacción más consciente con la tecnología.

De manera semejante, investigaciones recientes que combinan alfabetización en IA y aprendizaje autorregulado describen prácticas en las que los estudiantes planifican el uso de la herramienta, iteran sobre sus resultados y evalúan posteriormente tanto el producto obtenido como el proceso de colaboración con la IA. Estas dinámicas sugieren que la autonomía se fortalece cuando el estudiante conserva la capacidad de dirigir la interacción y establecer criterios propios de evaluación.

Así, los resultados permiten diferenciar entre autonomía asistida por IA y dependencia de la IA. En la primera, la tecnología amplía las posibilidades del estudiante sin sustituir sus decisiones; en la segunda, las decisiones

relevantes comienzan a trasladarse progresivamente hacia el sistema.

Dependencia, descarga cognitiva y pérdida potencial de agencia

La cuarta categoría concentró los hallazgos relacionados con los posibles efectos negativos de una utilización excesivamente dependiente de la inteligencia artificial. Los estudios revisados señalaron que la disponibilidad inmediata de respuestas puede reducir determinados esfuerzos asociados con búsqueda, elaboración, revisión y resolución autónoma de problemas.

Uno de los resultados más relevantes fue aportado por Darvishi et al. (2024), quienes analizaron el impacto de la asistencia basada en IA sobre la agencia estudiantil mediante un experimento realizado con 1.625 estudiantes de diez cursos. Los resultados mostraron que los participantes tendían a depender de la asistencia proporcionada por la IA en lugar de aprender progresivamente de ella. Cuando dicha asistencia era retirada, las estrategias de autorregulación podían compensar parcialmente su ausencia, aunque no alcanzaban el mismo nivel de eficacia inmediata.

Este hallazgo resulta especialmente relevante porque permite diferenciar entre rendimiento asistido y aprendizaje autónomo. Un estudiante puede producir mejores resultados mientras dispone de una herramienta de IA sin necesariamente haber desarrollado las capacidades que le permitirían reproducir ese desempeño de manera independiente.

La evidencia también mostró que la denominada descarga cognitiva puede tener una función beneficiosa cuando libera recursos mentales para tareas de mayor complejidad; sin embargo, adquiere un carácter problemático cuando implica delegar sistemáticamente operaciones necesarias para el aprendizaje. Entre estas operaciones se encuentran la búsqueda de información, la construcción de argumentos, la comparación de perspectivas, la elaboración de conclusiones y la evaluación de evidencias.

Los resultados, por tanto, no permiten afirmar que el uso de IA produzca necesariamente una pérdida de pensamiento crítico o autonomía. Más bien muestran que el riesgo aumenta cuando la interacción se caracteriza por una aceptación acrítica de las respuestas, escasa supervisión metacognitiva y ausencia de mecanismos de verificación.

Síntesis de los principales hallazgos

La comparación de los estudios permitió identificar un patrón común: la variable determinante no es únicamente cuánto utiliza el estudiante la inteligencia artificial, sino qué tipo de actividad cognitiva conserva durante esa interacción.

La Tabla 2 sintetiza las principales tendencias identificadas en la revisión.

Tabla 2. Síntesis temática de los hallazgos sobre IA generativa, pensamiento crítico y autonomía estudiantil

Eje de análisis	Principales hallazgos	Potencial educativo	Riesgo identificado
IA como apoyo cognitivo	Facilita explicaciones, retroalimentación, generación de ideas y resolución guiada de problemas	Andamiaje, personalización y acceso rápido a apoyo académico	Sustitución de procesos que el estudiante debería desarrollar
Pensamiento crítico	La calidad del aprendizaje depende de contrastar, verificar y cuestionar las respuestas	Evaluación de evidencias, argumentación y detección de errores	Aceptación de información plausible sin comprobación
Metacognición y autorregulación	Los apoyos metacognitivos favorecen planificación, monitoreo y evaluación del aprendizaje	Mayor control sobre el proceso y uso consciente de IA	Uso pasivo de la herramienta cuando falta orientación
Autonomía y agencia	La IA puede ampliar la capacidad de actuación del estudiante si esta conserva las decisiones centrales	Autonomía asistida y aprendizaje autodirigido	Transferencia progresiva de decisiones hacia la herramienta
Dependencia y descarga cognitiva	La automatización puede reducir esfuerzo en determinadas tareas	Liberación de recursos para actividades de mayor complejidad	Dependencia, aprendizaje superficial y menor ejercicio del juicio propio

Fuente. Elaboración de los autores a partir de los estudios incluidos en la revisión sistemática.

Por ende, la evidencia permitió identificar dos trayectorias claramente diferenciadas. La primera se produce cuando la IA funciona como un recurso para ampliar el razonamiento: el estudiante formula preguntas, analiza la respuesta obtenida, contrasta información, identifica posibles errores, reformula sus ideas y finalmente

construye una conclusión propia. En este escenario, la interacción con IA puede contribuir al pensamiento crítico, la metacognición y la autonomía.

La segunda trayectoria aparece cuando la respuesta generada es asumida como producto final. En este caso disminuyen las oportunidades para realizar procesos de búsqueda, contraste, argumentación y evaluación, incrementándose el riesgo de dependencia y delegación cognitiva. Los resultados de Darvishi et al. (2024) constituyen una evidencia particularmente significativa de esta tensión, al mostrar que una asistencia eficaz durante la actividad no implica necesariamente que el estudiante haya aprendido a desempeñarse de manera autónoma.

De esta forma, los resultados sugieren que la frontera entre apoyo y dependencia no se encuentra en la utilización de la inteligencia artificial en sí misma, sino en el grado de control cognitivo que conserva el estudiante durante la interacción. La IA puede proporcionar respuestas, alternativas y retroalimentación; sin embargo, la selección, valoración y decisión final continúan siendo procesos esenciales para transformar dicha información en conocimiento y, posteriormente, en juicio propio.

DISCUSIÓN

Los resultados de esta revisión permiten sostener que la relación entre inteligencia artificial generativa, pensamiento crítico y autonomía estudiantil es más compleja de lo que suele plantearse en discursos polarizados sobre el uso educativo de la IA. La evidencia analizada no respalda una visión determinista según la cual estas herramientas mejoran automáticamente el aprendizaje, ni tampoco una interpretación que las considere necesariamente perjudiciales. Más bien, los hallazgos muestran que su efecto depende del tipo de interacción que establezca el estudiante, del diseño pedagógico que acompañe su uso y, especialmente, del grado de control cognitivo que conserve durante el proceso.

En este sentido, Chan y Hu (2023) evidenciaron que los estudiantes reconocen beneficios claros en la utilización de inteligencia artificial generativa, entre ellos la rapidez de respuesta, el apoyo personalizado y la facilidad para acceder a explicaciones. Sin embargo, también identificaron preocupaciones relacionadas con la precisión, la ética y la confiabilidad de los contenidos generados. Este doble reconocimiento resulta especialmente relevante porque confirma que la incorporación de la IA no elimina la necesidad de formar capacidades de evaluación; por el contrario, hace más urgente desarrollar criterios para decidir cuándo una respuesta es útil, cuándo debe verificarse y cuándo debe ser descartada.

Desde esta perspectiva, el pensamiento crítico adquiere una función distinta a la que tradicionalmente ha tenido en el aula. Larson et al. (2024) plantean que la expansión de la IA generativa obliga a reconsiderar el papel del pensamiento crítico en los procesos formativos. Ya no basta con que el estudiante sea capaz de construir un argumento propio desde cero; también debe ser competente para examinar argumentos producidos por sistemas artificiales, detectar vacíos, reconocer supuestos y establecer si una respuesta aparentemente bien formulada posee realmente suficiente respaldo. Esto transforma al pensamiento crítico en una especie de filtro epistemológico entre la producción automatizada de información y la construcción de conocimiento.

Los hallazgos obtenidos también coinciden con Essel et al. (2024), quienes analizaron los efectos de ChatGPT sobre habilidades cognitivas como el pensamiento crítico, creativo y reflexivo. Sus resultados muestran que la interacción con sistemas generativos puede estimular capacidades cognitivas cuando la actividad obliga al estudiante a participar activamente. Esto permite interpretar que el problema no está en utilizar una respuesta generada por IA, sino en el modo en que esa respuesta se incorpora al proceso de aprendizaje. Una respuesta utilizada como punto final puede reducir el esfuerzo cognitivo; la misma respuesta, utilizada como objeto de análisis, comparación o cuestionamiento, puede convertirse en un recurso para profundizar el aprendizaje.

Esta diferencia es fundamental para comprender el concepto de autonomía. Chiu (2024) sostiene que el uso de inteligencia artificial puede favorecer procesos de aprendizaje autorregulado cuando la herramienta se integra con mecanismos que promueven motivación, reflexión y control del propio aprendizaje. De manera similar, Ng et al. (2024) encontraron que ChatGPT puede contribuir al aprendizaje autorregulado cuando los estudiantes reciben orientación adecuada sobre cómo utilizarlo. Ambos trabajos permiten cuestionar una idea frecuente: que utilizar IA reduce necesariamente la independencia del estudiante. En realidad, la autonomía no depende de prescindir de herramientas, sino de conservar la capacidad de dirigirlas.

Bajo esta lógica, un estudiante autónomo no es aquel que trabaja sin apoyo tecnológico, sino aquel que puede decidir qué preguntar, cómo interpretar la respuesta, cuándo comprobarla y qué parte de ella resulta útil para alcanzar sus objetivos. Esta interpretación es coherente con Yang et al. (2024), quienes diferencian entre aproximaciones superficiales y profundas al aprendizaje con IA generativa. Su análisis muestra que la agencia estudiantil desempeña un papel central en la manera en que las herramientas generativas se integran al proceso educativo. Cuando el estudiante mantiene un papel activo, la IA puede apoyar formas de aprendizaje más profundas; cuando asume un papel pasivo, existe mayor riesgo de utilización superficial.

Uno de los hallazgos más importantes de la revisión es precisamente esta tensión entre asistencia y dependencia. Darvishi et al. (2024) mostraron que los estudiantes pueden llegar a depender de la asistencia de la IA en lugar de aprender progresivamente de ella. Este resultado obliga a diferenciar entre desempeño inmediato y aprendizaje real. La obtención de una respuesta correcta con apoyo tecnológico no demuestra necesariamente que el estudiante haya desarrollado las competencias necesarias para resolver un problema similar de forma independiente.

Esta distinción tiene implicaciones educativas importantes. La eficiencia suele considerarse una ventaja de la inteligencia artificial, pero desde una perspectiva pedagógica no toda reducción del esfuerzo es necesariamente beneficiosa. Determinados esfuerzos cognitivos, como formular una hipótesis, buscar evidencias, comparar alternativas o construir una conclusión, forman parte del proceso mediante el cual se aprende. Cuando estas tareas son delegadas completamente a la IA, puede producirse una aparente mejora del producto final mientras disminuye la participación intelectual del estudiante.

En este punto resulta especialmente útil el concepto de descarga cognitiva. Iqbal et al. (2025) analizaron la relación entre el uso de herramientas generativas, metacognición compartida y cognitive offloading. Sus resultados sugieren que la descarga cognitiva no debe considerarse intrínsecamente negativa. Delegar tareas rutinarias puede liberar recursos mentales para procesos más complejos. El problema aparece cuando la descarga deja de ser estratégica y comienza a incluir funciones esenciales para el aprendizaje, como razonar, interpretar, contrastar o decidir.

La frontera entre una descarga cognitiva útil y una dependencia problemática parece encontrarse precisamente en la metacognición. Xu et al. (2025) destacan que el apoyo metacognitivo desempeña un papel crítico en entornos de aprendizaje con inteligencia artificial generativa. Cuando el estudiante planifica su interacción, monitorea su comprensión y evalúa la calidad de las respuestas recibidas, la IA puede incorporarse dentro de un proceso de aprendizaje autorregulado. Sin estos mecanismos, existe mayor probabilidad de que la herramienta sea utilizada de forma reactiva y poco reflexiva.

Esta cuestión se relaciona también con los hallazgos de Urban et al. (2025), quienes analizaron las creencias epistémicas y la precisión metacognitiva de estudiantes al integrar contenido generado por ChatGPT en escritura académica. Su trabajo resulta especialmente significativo porque muestra que disponer de una advertencia sobre posibles errores de la IA no garantiza que los estudiantes puedan identificarlos adecuadamente. Esto implica que la simple recomendación de “verificar la información” es insuficiente si el estudiante no posee criterios para realizar dicha verificación.

A partir de esta evidencia, la alfabetización en IA debería superar una visión centrada exclusivamente en competencias instrumentales. Walter (2024) propone integrar alfabetización en IA, ingeniería de prompts y pensamiento crítico como competencias esenciales de la educación contemporánea. Esta propuesta resulta coherente con los hallazgos de la presente revisión, aunque conviene ampliar todavía más el concepto: saber utilizar IA implica también saber desconfiar de ella de manera razonada, reconocer sus límites y asumir responsabilidad sobre el contenido finalmente utilizado.

En la misma línea, Tzirides et al. (2024) plantean la necesidad de combinar inteligencia humana y artificial dentro de procesos educativos que fortalezcan la alfabetización en IA. Este enfoque resulta particularmente útil porque evita presentar la relación humano–máquina como una competencia entre ambos. La cuestión no es decidir si debe pensar el estudiante o la IA, sino establecer qué funciones conviene delegar y cuáles deben permanecer bajo control humano. El juicio, la valoración ética, la contextualización y la decisión final constituyen precisamente los ámbitos en los que la participación del estudiante sigue siendo irremplazable.

Los resultados de esta revisión permiten, por tanto, proponer una interpretación basada en dos trayectorias posibles. La primera puede definirse como una trayectoria de ampliación cognitiva: el estudiante utiliza la

respuesta de la IA como insumo, la contrasta, identifica limitaciones, reformula preguntas y finalmente produce una conclusión propia. En esta trayectoria, la IA funciona como un interlocutor que amplía las posibilidades de aprendizaje sin sustituir la toma de decisiones.

La segunda trayectoria corresponde a una forma de sustitución cognitiva: el estudiante obtiene una respuesta, la acepta con escasa evaluación y la incorpora directamente a la tarea. Cuando este comportamiento se repite, funciones como buscar, interpretar, argumentar y decidir pueden ser progresivamente externalizadas. En este escenario, la IA deja de actuar como apoyo y comienza a ocupar el lugar del razonamiento que debería desarrollar el estudiante.

Esta interpretación coincide con Farrokhnia et al. (2024), quienes advierten que las ventajas educativas de ChatGPT conviven con riesgos relacionados con dependencia, integridad académica y debilitamiento de habilidades cognitivas de orden superior. Sin embargo, el análisis realizado permite añadir un matiz importante: estos riesgos no deberían abordarse únicamente mediante restricciones o prohibiciones. Si el problema central es la aceptación acrítica, la respuesta pedagógica debe consistir en diseñar tareas que obliguen al estudiante a cuestionar, justificar y verificar.

En este sentido, la educación superior podría desplazarse progresivamente desde actividades centradas en la producción de respuestas hacia actividades centradas en la evaluación de respuestas. Pedir a un estudiante que genere un resumen, explique un concepto o redacte un argumento son tareas que la IA puede realizar con facilidad. En cambio, solicitarle que identifique los errores de una respuesta generada, contraste fuentes, determine qué argumento resulta más sólido o justifique por qué rechaza una conclusión exige un nivel mayor de participación intelectual.

Esta transformación tiene consecuencias directas para la evaluación. Si la inteligencia artificial puede producir productos académicos plausibles, evaluar únicamente el producto final ofrece información cada vez más limitada sobre el aprendizaje. Resulta necesario valorar también el proceso seguido por el estudiante: cómo formuló sus preguntas, qué evidencias utilizó, qué modificaciones realizó, qué decisiones tomó y cómo justificó esas decisiones. De esta manera, la evaluación puede recuperar una función formativa y contribuir al desarrollo de autonomía.

Los hallazgos permiten afirmar que el verdadero desafío educativo no es evitar que los estudiantes utilicen inteligencia artificial, sino impedir que deleguen en ella la responsabilidad de pensar. La autonomía no se pierde cuando se utiliza una herramienta externa; se debilita cuando el estudiante deja de establecer criterios propios para decidir qué aceptar, qué cuestionar y qué rechazar.

Por ello, la principal contribución de esta revisión consiste en desplazar el debate desde la pregunta “¿deben los estudiantes utilizar inteligencia artificial?” hacia una cuestión pedagógicamente más relevante: “¿qué tipo de pensamiento debe conservar y ejercer el estudiante mientras utiliza inteligencia artificial?”. A partir de la evidencia analizada, la respuesta parece encontrarse en el fortalecimiento conjunto del pensamiento crítico, la metacognición y la agencia estudiantil. Solo cuando estos procesos permanecen activos es posible avanzar verdaderamente de la respuesta proporcionada por la IA al juicio propio del estudiante.

CONCLUSIONES

La revisión realizada permite concluir que la inteligencia artificial generativa está transformando de manera profunda las formas en que los estudiantes acceden, procesan y producen información. Sin embargo, su incorporación en los procesos educativos no garantiza por sí sola mejores aprendizajes ni una mayor autonomía. El efecto formativo de estas herramientas depende, en gran medida, del tipo de interacción que el estudiante establezca con ellas y del nivel de control cognitivo que conserve durante su utilización.

Los hallazgos muestran que la IA puede desempeñar una función positiva cuando actúa como apoyo para explorar ideas, recibir retroalimentación, contrastar perspectivas y profundizar la comprensión. En estas condiciones, su uso puede favorecer procesos de autorregulación, metacognición y aprendizaje autónomo. No obstante, cuando la respuesta generada es asumida como producto final y se reduce la participación del estudiante en tareas de análisis, verificación, argumentación y toma de decisiones, aumenta el riesgo de dependencia y delegación excesiva del esfuerzo cognitivo.

En este contexto, el pensamiento crítico se configura como una competencia esencial para el uso educativo de la inteligencia artificial. Ya no resulta suficiente que el estudiante sepa formular preguntas o utilizar

técnicamente una herramienta; necesita también evaluar la calidad de las respuestas, reconocer posibles errores o sesgos, contrastar fuentes y justificar las decisiones que adopta a partir de la información recibida. La autonomía, por tanto, no debe entenderse como ausencia de apoyo tecnológico, sino como la capacidad de mantener el control intelectual sobre el propio proceso de aprendizaje.

Uno de los principales aportes de este estudio consiste en identificar dos trayectorias posibles en la interacción entre estudiante e IA. La primera corresponde a una trayectoria de ampliación cognitiva, en la que la respuesta generada es sometida a análisis, contraste y reflexión hasta convertirse en una conclusión construida por el propio estudiante. La segunda representa una trayectoria de sustitución cognitiva, caracterizada por la aceptación acrítica de las respuestas, la descarga excesiva de procesos mentales y una progresiva pérdida de agencia. Esta distinción permite comprender que el problema no reside en la tecnología en sí misma, sino en la manera en que se integra dentro del proceso educativo.

A partir de estos resultados, las instituciones de educación superior deberían avanzar hacia estrategias pedagógicas que enseñen a trabajar críticamente con inteligencia artificial. Esto implica diseñar actividades en las que los estudiantes deban verificar, comparar, argumentar, corregir y justificar las respuestas generadas, así como incorporar mecanismos de reflexión metacognitiva que permitan hacer visible el proceso seguido para llegar a una conclusión.

También resulta necesario reconsiderar los modelos tradicionales de evaluación. En un contexto donde la IA puede producir textos y respuestas académicamente plausibles, la valoración del aprendizaje no debería centrarse exclusivamente en el producto final. Evaluar el razonamiento, las fuentes consultadas, las decisiones tomadas y la capacidad del estudiante para explicar y defender sus conclusiones puede ofrecer una aproximación más auténtica al desarrollo del pensamiento crítico y de la autonomía.

Como limitación, la revisión se concentró principalmente en literatura reciente, debido al carácter emergente de la inteligencia artificial generativa en educación. Además, la heterogeneidad metodológica de los estudios incluidos dificulta establecer relaciones causales generalizables entre uso de IA, pensamiento crítico y autonomía. Futuras investigaciones podrían profundizar mediante estudios longitudinales, experimentales y comparativos que permitan determinar cómo evolucionan estas competencias cuando la utilización de IA se mantiene durante periodos prolongados.

En síntesis, la inteligencia artificial puede ampliar las capacidades del estudiante, pero no debería reemplazar aquello que constituye el núcleo del aprendizaje: comprender, cuestionar, decidir y construir una posición propia. El desafío educativo contemporáneo no consiste únicamente en enseñar a obtener mejores respuestas de la IA, sino en formar estudiantes capaces de ir más allá de esas respuestas y convertirlas, mediante reflexión y criterio, en conocimiento verdaderamente propio.

A partir de los hallazgos de esta revisión, se recomienda que la incorporación de inteligencia artificial generativa en la educación superior se acompañe de estrategias pedagógicas orientadas a preservar y fortalecer el papel activo del estudiante en la construcción del conocimiento. El desafío no consiste únicamente en enseñar a utilizar estas herramientas, sino en formar estudiantes capaces de cuestionar, contrastar y justificar la información que reciben de ellas.

En primer lugar, resulta conveniente diseñar actividades en las que la IA funcione como punto de partida y no como producto final. Las tareas académicas pueden orientarse a comparar respuestas generadas, identificar errores, contrastar fuentes, argumentar desacuerdos y reformular conclusiones. De esta manera, el estudiante mantiene una participación intelectual activa y desarrolla criterios propios de valoración.

También se recomienda fortalecer explícitamente la alfabetización en inteligencia artificial, incorporando no solo competencias técnicas relacionadas con el uso de herramientas y elaboración de instrucciones, sino también conocimientos sobre sus limitaciones, posibles sesgos, errores y condiciones de confiabilidad. Comprender cómo y por qué una IA puede equivocarse favorece una relación más crítica y responsable con la tecnología.

Asimismo, es necesario integrar estrategias de autorregulación y metacognición en las experiencias de aprendizaje mediadas por IA. El estudiante debería ser capaz de reconocer para qué utiliza la herramienta, qué parte de la tarea está delegando, qué ha aprendido durante el proceso y si pudiera defender la respuesta obtenida sin depender exclusivamente del sistema.

En materia de evaluación, se recomienda reducir la dependencia de actividades centradas únicamente en el producto final. Resulta pertinente valorar también el proceso de razonamiento, las decisiones adoptadas, las

fuentes consultadas, las modificaciones realizadas a las respuestas generadas y la capacidad del estudiante para explicar y defender sus conclusiones. Este enfoque permite evaluar con mayor autenticidad el pensamiento crítico y la autonomía.

Del mismo modo, el profesorado requiere espacios de formación que le permitan diseñar experiencias educativas en las que la IA sea utilizada como recurso de apoyo y no como sustituto del razonamiento. La formación docente debería incluir estrategias para formular actividades de mayor complejidad cognitiva, identificar usos superficiales de la IA y promover dinámicas de verificación, reflexión y argumentación.

En el ámbito institucional, se recomienda desarrollar orientaciones claras sobre el uso académico de inteligencia artificial, evitando tanto la prohibición generalizada como la aceptación irrestricta. Las políticas deberían priorizar principios de transparencia, responsabilidad, integridad académica y autonomía, diferenciando entre usos legítimos de apoyo y formas de utilización que sustituyan indebidamente procesos esenciales de aprendizaje.

Finalmente, futuras investigaciones deberían examinar los efectos del uso prolongado de IA generativa sobre el pensamiento crítico y la autonomía mediante diseños longitudinales, experimentales y comparativos. También sería pertinente estudiar diferencias entre disciplinas, niveles educativos y tipos de tareas, así como identificar qué estrategias pedagógicas permiten transformar la asistencia de la IA en una oportunidad real para fortalecer el juicio propio.

REFERENCIAS BIBLIOGRÁFICAS

Baig, M. I., & Yadegaridehkordi, E. (2024). ChatGPT in the higher education: A systematic literature review and research challenges. International Journal of Educational Research, 127, 102411. DOI: 10.1016/j.ijer.2024.102411.

Bosch, C., & Kruger, D. (2024). AI chatbots as Open Educational Resources: Enhancing student agency and self-directed learning. Italian Journal of Educational Technology. DOI: 10.17471/2499-4324/1326.

Camastra, F. D., & González Vallejo, R. (2025). Inteligencia artificial, sostenibilidad e impacto ambiental. Un estudio narrativo y bibliométrico. Región Científica, 4(1), 2025355. https://doi.org/10.58763/rc2025355

Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. International Journal of Educational Technology in Higher Education, 20, 38. DOI: 10.1186/s41239-023-00408-3.

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. International Journal of Educational Technology in Higher Education, 20, 43. DOI: 10.1186/s41239-023-00411-8.

Chiu, T. K. F. (2024). A classification tool to foster self-regulated learning with generative artificial intelligence by applying self-determination theory: A case of ChatGPT. Educational Technology Research and Development, 72, 2401–2416. DOI: 10.1007/s11423-024-10366-w.

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. Innovations in Education and Teaching International, 61(2), 228–239. DOI: 10.1080/14703297.2023.2190148.

Darvishi, A., Khosravi, H., Sadiq, S., Gašević, D., & Siemens, G. (2024). Impact of AI assistance on student agency. Computers & Education, 210, 104967. DOI: 10.1016/j.compedu.2023.104967.

Essel, H. B., Vlachopoulos, D., Essuman, A. B., & Amankwa, J. O. (2024). ChatGPT effects on cognitive skills of undergraduate students: Receiving instant responses from AI-based conversational large language models (LLMs). Computers and Education: Artificial Intelligence, 6, 100198. DOI: 10.1016/j.caeai.2023.100198.

Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2024). A SWOT analysis of ChatGPT: Implications for educational practice and research. Innovations in Education and Teaching International, 61(3), 460–474. DOI: 10.1080/14703297.2023.2195846.

Iqbal, J., Hashmi, Z. F., Asghar, M. Z., & Abid, M. N. (2025). Generative AI tool use enhances academic achievement in sustainable education through shared metacognition and cognitive offloading among preservice teachers. Scientific Reports, 15, 16610. DOI: 10.1038/s41598-025-01676-x.

Larson, B. Z., Moser, C., Caza, A., Muehlfeld, K., & Colombo, L. A. (2024). Critical thinking in the age of generative AI. Academy of Management Learning & Education. DOI: 10.5465/amle.2024.0338.

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. Education Sciences, 13(4), 410. DOI: 10.3390/educsci13040410.

Martínez-Peláez, R., Mena, L. J., Toral-Cruz, H., Ochoa-Brust, A., González Potes, A., Flores, V., Ostos, R., Ramírez Pacheco, J. C., Félix, R. A., & Félix, V. G. (2025). Can large language models foster critical thinking, teamwork, and problem-solving skills in higher education? A literature review. Systems, 13(11), 1013. DOI: 10.3390/systems13111013.

Ng, D. T. K., Tan, C. W., & Leung, J. K. L. (2024). Empowering student self-regulated learning and science education through ChatGPT: A pioneering pilot study. British Journal of Educational Technology, 55(4), 1328–1353. DOI: 10.1111/bjet.13454.

Padilla Hernández, S. G. (2026). Confianza del cliente en la inteligencia artificial aplicada a servicios bancarios: tendencias y agenda de investigación. Región Científica, 5(1), 2026520. https://doi.org/10.58763/rc2026520

Padilla Hernández, S. G. (2024). Inteligencia Artificial en los servicios bancarios. Una revisión bibliométrica. Región Científica, 3(2), 2024335. https://doi.org/10.58763/rc2024335

Santana González, Y. (2025). La inteligencia artificial DeepSeek en la educación médica cubana: oportunidad para liderar la innovación global. Revista Cubana de Medicina, 64, e5061. https://revmedicina.sld.cu/index.php/med/article/view/5061

Sardiñas Padilla, E., & Valdés García, K. (2025). Incorporación de la inteligencia artificial a la educación cubana. Ventajas y limitaciones. Región Científica, 4(1), 2025372. https://doi.org/10.58763/rc2025372

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. DOI: 10.1016/j.caeo.2024.100169.

Suryati, N., Laksmi, E. D., & Marzuki. (2023). Critical thinking in the AI era: An exploration of EFL students' perceptions, benefits, and limitations. Cogent Education. DOI: 10.1080/2331186X.2023.2290342.

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, 15. DOI: 10.1186/s40561-023-00237-x.

Trombetta Silvero, C. E., Patiño Sotelo, A., Basili Carreras, J. L., & Pintos Coronel, M. R. (2026). El rol de la inteligencia artificial en la formación de ingenieros agrónomos: oportunidades y desafíos. Región Científica, 5(1), 2026595. https://doi.org/10.58763/rc2026595

Tzirides, A. O., Zapata, G., Kastania, N. P., Saini, A. K., Castro, V., Ismael, S. A., You, Y.-L., Afonso dos Santos, T., Searsmith, D., O'Brien, C., Cope, B., & Kalantzis, M. (2024). Combining human and artificial intelligence for enhanced AI literacy in higher education. Computers and Education Open, 6, 100184. DOI: 10.1016/j.caeo.2024.100184.

Urban, M., Brom, C., Lukavský, J., Děchtěrenko, F., Hein, V., Svacha, F., Kmoníčková, P., & Urban, K. (2025). "ChatGPT can make mistakes. Check important info." Epistemic beliefs and metacognitive accuracy in students' integration of ChatGPT content into academic writing. British Journal of Educational Technology, 56, 1897–1918. DOI: 10.1111/bjet.13591.

Walter, Y. (2024). Embracing the future of Artificial Intelligence in the classroom: The relevance of AI literacy, prompt engineering, and critical thinking in modern education. International Journal of Educational Technology in Higher Education, 21, 15. DOI: 10.1186/s41239-024-00448-3.

Xia, L., Shen, K., Sun, H., An, X., et al. (2025). Developing and validating the student learning agency scale in generative artificial intelligence (AI)-supported contexts. Education and Information Technologies, 30, 13999–14021. DOI: 10.1007/s10639-024-13137-5.

Xu, X., Qiao, L., Cheng, N., Liu, H., & Zhao, W. (2025). Enhancing self-regulated learning and learning experience in generative AI environments: The critical role of metacognitive support. British Journal of Educational Technology, 56(5), 1842–1863. DOI: 10.1111/bjet.13599.

Yang, Y., Luo, J., Yang, M., Yang, R., & Chen, J. (2024). From surface to deep learning approaches with Generative AI in higher education: An analytical framework of student agency. Studies in Higher Education, 49(5), 817–830. DOI: 10.1080/03075079.2024.2327003.

Yao, Y., Sun, Y., Zhu, S., & Zhu, X. (2025). A qualitative inquiry into metacognitive strategies of postgraduate students in employing ChatGPT for English academic writing. European Journal of Education, 60(1), e12824. DOI: 10.1111/ejed.12824.

Zapata Muriel, F. A., Montoya Zapata, S., & Montoya-Zapata, D. (2024). Dilemas éticos planteados por el auge de la inteligencia artificial: una mirada desde el transhumanismo. Región Científica, 3(1), 2024225. https://doi.org/10.58763/rc2024225

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Declaración de responsabilidad de autoría

Los autores del manuscrito señalado, DECLARAMOS que hemos contribuido directamente a su contenido

intelectual, así como a la génesis y análisis de sus datos; por lo cual, estamos en condiciones de hacernos públicamente responsable de él y aceptamos que sus nombres figuren en la lista de autores en el orden indicado. Además, hemos cumplido los requisitos éticos de la publicación mencionada, habiendo consultado la Declaración de Ética y mala praxis en la publicación.

Los cuatro autores participamos en el proceso de revisión de literatura, análisis de los datos obtenidos y redacción del artículo. Juan de Jesús Alvarado Ortiz, Oralia Mercedes Moguel Morales, Francisco Jesús Vieyra González, Onelia Hernández Torres