Editorial Neogranadina

Revista Prolegómenos 2025 julio-diciembre Vol. 28(56) ISSN: 0121-182x · e-ISSN: 1909-7727 ■ pp. 31-47

DOI: https://doi.org/10.18359/prole.7515

Inteligencia artificial en la formación jurídica: análisis de su integración a la investigación formativa*

Gabriel Ravelo-Franco[a]

Resumen: El estudio examina la integración de la inteligencia artificial (IA), especialmente mediante la herramienta ChatGPT-4o, en la formación jurídica, con el fin de evaluar su impacto sobre la integridad académica, las habilidades analíticas y el desarrollo ético de estudiantes de derecho. Basado en una revisión de treinta estudios recientes (2022-2024), se realizó una experimentación práctica generando resúmenes críticos de sentencias del Tribunal Constitucional del Perú con ChatGPT-4o, sometiéndolos a evaluaciones de originalidad mediante Turnitin y ChatGPTZero. Los resultados evidencian que la IA puede mejorar la alfabetización digital, ofrecer aprendizaje personalizado y proporcionar retroalimentación precisa y coherente similar al estilo humano. Sin embargo, también se identifican riesgos relacionados con la integridad académica, lo que incluye plagio, suplantación de autoría y uso inapropiado en tareas académicas. Se concluye la necesidad de implementar controles efectivos, políticas claras y marcos éticos idóneos, además de capacitar a estudiantes y docentes en competencias éticas y técnicas específicas para maximizar los beneficios educativos de la IA y mitigar sus riesgos éticos y académicos.

Palabras clave: educación jurídica; integridad académica; sesgos algorítmicos; inteligencia artificial; formación jurídica

Recibido: 13/08/2024/ **Aceptado:** 05/05/2025 **Disponible en línea:** 07/10/2025

Cómo citar: Ravelo-Franco, G. (2025). Inteligencia artificial en la formación jurídica: Análisis de su integración a la investigación formativa. *Prolegómenos*, 28(56), 31–47. https://doi.org/10.18359/prole.75155

* Artículo de investigación.

a Magíster en derecho penal. Profesor de Facultad de Derecho, Universidad Continental, Huancayo, Perú. Correo electrónico: gravelo@continental.edu.pe; ORCID: https://orcid.org/0000-0003-0212-312X

Artificial Intelligence in Legal Education: Analysis of its Integration into Formative Research

Abstract: The study examines the integration of artificial intelligence (AI), particularly through the use of the ChatGPT-4o tool, into legal education, aiming to assess its impact on academic integrity, analytical skills, and the ethical development of law students. Based on a review of 30 recent studies (2022-2024), a practical experiment was conducted using ChatGPT-4o to generate critical summaries of rulings issued by the Constitutional Court of Peru. These summaries were subsequently subjected to originality assessments using Turnitin and ChatGPTZero. Findings indicate that AI can significantly enhance digital literacy, facilitate personalized learning, and provide accurate and coherent feedback similar to human-generated content. However, notable risks associated with academic integrity were also identified, including plagiarism, authorship fraud, and inappropriate reliance on AI for academic tasks. This paper concludes by emphasizing the urgent need for effective controls, clear institutional policies, and robust ethical frameworks. Furthermore, it highlights the importance of training both students and educators in specific ethical and technical competencies to maximize the educational benefits of AI while effectively mitigating its ethical and academic risks.

Key-words: Law Education; Academic Integrity; Algorithmic Biases; Artificial Intelligence; Legal Education

Inteligência artificial na formação jurídica: análise de sua integração à pesquisa formativa

Resumo: O estudo examina a integração da inteligência artificial (IA), especialmente por meio da ferramenta ChatGPT-4o, na formação jurídica, com o objetivo de avaliar seu impacto sobre a integridade acadêmica, as habilidades analíticas e o desenvolvimento ético de estudantes de Direito. Com base na revisão de trinta estudos recentes (2022-2024), foi realizada uma experimentação prática, gerando resumos críticos de sentenças do Tribunal Constitucional do Peru com o ChatGPT-4o e submetendo-os a avaliações de originalidade por meio do Turnitin e do ChatGPTZero. Os resultados mostram que a IA pode melhorar a alfabetização digital, oferecer aprendizado personalizado e fornecer feedback preciso e coerente, similar ao estilo humano. No entanto, também foram identificados riscos relacionados à integridade acadêmica, incluindo plágio, falsificação de autoria e uso inadequado em tarefas acadêmicas. Conclui-se que é necessário implementar controles eficazes, políticas claras e marcos éticos adequados, além de capacitar estudantes e docentes em competências éticas e técnicas específicas para maximizar os benefícios educacionais da IA e mitigar seus riscos éticos e acadêmicos.

Palavras-chave: educação jurídica; integridade acadêmica; vieses algorítmicos; inteligência artificial; formação jurídica

Introducción

La IA ha comenzado a transformar diversos ámbitos, incluida la educación, lo que ha generado interés por parte de la academia. Esto ha dado lugar a estudios que se enfocan en el provecho que se puede obtener de tales herramientas, mientras que otros se centran en los riesgos y desafíos derivados de las mismas.

En el primer grupo, Galvis Martínez y Esquivel Zambrano (2022) plantean que los objetivos de la IA están centrados en el desarrollo de métodos para igualar la inteligencia humana, lo que abarca aspectos como la conciencia y el pensamiento. Agregan que la integración de la IA en la robótica apunta a mejorar la calidad de vida humana, ofreciendo alternativas de bienestar en diversos ámbitos cotidianos. Asimismo, señalan que, aunque hay varias opiniones al respecto, es evidente que su incorporación proporciona beneficios para el desarrollo humano.

Dentro de la misma línea, Gallent Torres *et al*. (2023) se enfocaron en la aplicación de la IA en el ámbito educativo, enfatizando en la importancia de reconocer que la evolución es parte integral del progreso y beneficia al individuo; sin embargo, añaden que esta circunstancia no libera de la responsabilidad de analizar críticamente las posibles implicaciones y consecuencias de la incorporación de estas nuevas tecnologías al proceso educativo.

En el grupo de trabajos que identifican riesgos, se encuentra el de Navarro-Dolmestch (2023), quien encuentra un claro ejemplo de los riesgos de la IA: la oportunidad que brinda ChatGPT para reemplazar al estudiante en sus evaluaciones, especialmente en aquellas que requieren escritos (p. 257). En una línea similar, Luna Martínez (2023) afirma que, como humanos, no podemos limitarnos a tener solo respuestas éticas superficiales, aceptar sin cuestionamientos el uso de la tecnología, verla como un medio invencible de control sobre la vida, o simplemente rechazarla y condenarnos a quedarnos tecnológicamente atrás. Para dicho autor, debemos enfrentarla de manera equilibrada, por lo que los educadores no deben limitar su rol a la mera implementación acrítica de tecnología, sino que deben fomentar un enfoque crítico basado en un entendimiento profundo de sus implicaciones. El objetivo, añade, es cultivar individuos reflexivos que utilicen la tecnología con responsabilidad social, solidaridad, principios democráticos, respeto por la dignidad humana y cuidado del medio ambiente, para equilibrar las intenciones de grandes capitales y gobiernos autoritarios que buscan establecer un control digital total.

Dentro del panorama expuesto, es posible reconocer una tendencia a afirmar la necesidad de equilibrar las ventajas provenientes de la IA en el proceso educativo frente a los inevitables riesgos que ella conlleva. En esa tendencia se encuentran las posturas de Norman-Acevedo (2023), quien afirma que la incorporación de tecnologías de IA en la enseñanza de tutores virtuales y profesores universitarios puede ser muy beneficiosa para el desarrollo y aprendizaje de los estudiantes. El autor también sostiene que se deben considerar tanto las ventajas como las posibles desventajas de estas tecnologías: aunque actualmente solo percibimos la superficie de su potencial, no podemos permitirnos esperar a comprenderlo completamente antes de actuar o, de lo contrario, nos quedaremos atrás.

De la revisión preliminar de la literatura se ha podido identificar una laguna de conocimiento, la cual versa sobre la aplicabilidad de las herramientas de IA en la formación de abogadas y abogados. Este proyecto busca así superar dicha laguna en la comprensión actual del papel que la IA puede desempeñar en la educación jurídica. Mediante un análisis crítico y evaluativo, aspira a determinar en qué medida herramientas como ChatGPT-4o pueden enriquecer, o dificultar, la formación de futuros abogados y garantizar que su integración en la enseñanza jurídica resulte beneficiosa y se realice de manera responsable desde el punto de vista ético.

Es por ello que este proyecto investiga la incorporación de la IA, con énfasis en ChatGPT-4o, en la formación de estudiantes de derecho, examinando su influencia en la integridad académica, los sesgos potenciales y las capacidades de análisis crítico. A través de una revisión de la literatura y de una experimentación exploratoria, sin participación de

sujetos humanos, este estudio pretende proporcionar perspectivas sobre la eficacia y los retos del uso de la IA en la formación jurídica.

Metodología

La revisión de la literatura aborda investigaciones sobre la aplicación de la IA en educación, publicadas desde 2022 hasta la actualidad en los repositorios Scielo y Redalyc. El primer filtro estuvo compuesto por las siguientes expresiones de búsqueda: "inteligencia artificial" AND "educación", y se utilizó el comando *site* para la búsqueda avanzada a través de Google (ejemplo: "inteligencia artificial" +"educación" site:scielo.cl/pdf). Se usó dicho comando debido a las limitaciones de funcionalidad de su buscador nativo (search.scielo.org). Como ejemplo, cuando se aplica dicho buscador, para el caso de la colección chilena arroja 4 resultados, mientras que con el comando *site* de Google se obtienen 402 resultados. Luego se filtraron los artículos por años de publicación (2022 a 2024) y se seleccionaron aquellos relevantes para la educación universitaria, a través de expresiones de búsqueda adicionales como: fraude académico, integridad académica, plagio y sesgos.

En total, se incluyeron en el estudio treinta trabajos académicos que corresponden a los criterios de inclusión y a partir de los cuales se ejecutó la revisión de la literatura. La comparación se realizó en función de dos ejes temáticos: integridad académica y sesgos de las IA.

Por otro lado, se realizó un requerimiento de generación de textos académicos. Los textos generados no fueron modificados, siendo enteramente redactados por ChatGPT-4o. Dichos textos fueron sometidos a detección de similitud a través de Turnitin y a detección de generación por IA a través de ChatGPTZero.

Inteligencia artificial e integridad académica

A partir de la revisión de la literatura seleccionada y de acuerdo a los criterios descritos en la sección de metodología, en la tabla 1 se sintetiza la postura de cada trabajo respecto de consideraciones relativas a la integridad académica en el contexto de la IA.

Tabla 1. Inteligencia artificial e integridad académica

Autor (año)	Resultados relativos
Galvis Martínez y Esquivel Zambrano (2022)	El Parlamento Europeo ha solicitado a la Comisión Europea que establezca un régimen de responsabilidad para los fabricantes de robots y desarrolladores de IA, el cual incluye la creación de un fondo económico para indemnizaciones por posibles daños. Además, se enfoca en la urgencia de un código de ética para los investigadores, para asegurar que los robots actúen de manera ética y beneficiosa para la humanidad.
Calvo (2022)	En 2018, la Comisión Europea adoptó los principios de la bioética como guía para la orientación de los investigadores en el documento *Draft Ethics Guidelines for Trustworthy AI*, adaptándolos al contexto de la IA y añadiendo a los cuatro principios definidos de beneficencia, autonomía, no-maleficencia y justicia un nuevo principio: explicabilidad.
Gallent Torres *et al*. (2023)	El impacto de la IA en la educación superior debe realizarse desde dos frentes: desde el estudio de dichas herramientas con énfasis en su aplicación para obtener mejores resultados y desde la reflexión respecto de sus implicancias y limitaciones éticas.
García Peñalvo *et al*. (2023)	Para frenar el uso fraudulento y poco ético de tecnologías, el profesorado debe concientizar a los estudiantes sobre la importancia de la honestidad académica, el pensamiento crítico y las consecuencias del engaño (Choi *et al*., 2023; Dwivedi *et al*., 2023). Además, debe ejercer un mayor liderazgo y fomentar una responsabilidad compartida entre gobierno, profesorado y estudiantes (Crawford *et al*., 2023; Lim *et al*., 2023).
Jiménez-García *et al*. (2023)	Es fundamental enfrentar los desafíos éticos, morales y de privacidad en la gestión de datos personales, como indican Renz y Vladova (2021) y Yuskovych-Zhukovska *et al*. (2022). Para ello, se requiere implementar políticas claras y medidas de protección que prevengan la discriminación o el daño a los estudiantes, para así aprovechar los beneficios de la IA sin comprometer la privacidad ni la ética. Asimismo, es necesario que tanto estudiantes como docentes desarrollen competencias adaptadas a la era de la IA, que incluyan la comprensión de sus aspectos éticos y sociales, como la privacidad y seguridad de los datos, así como su impacto en la sociedad (Arrieta *et al*., 2020; Unesco, 2021).

Autor (año)	Resultados relativos
Norman-Acevedo (2023)	La incorporación de tecnologías de IA en la enseñanza de tutores y profesores universitarios puede ser un recurso útil para potenciar el aprendizaje y el desarrollo de competencias en los estudiantes. No obstante, es necesario evaluar tanto las ventajas como las desventajas que esta tecnología conlleva.
Navarro-Dolmestch (2023)	La fiscalización institucional de conductas inapropiadas como el plagio y la suplantación de autoría pueden motivar intentos de evadir los controles, así como afectar los valores académicos de justicia y responsabilidad. Un conjunto de mecanismos de supervisión y control puede reprimir estas conductas, pero también puede generar riesgos para la viabilidad del modelo pedagógico.
Forero-Corba y Negre Bennasar (2023)	Los procesos educativos, así como las técnicas y herramientas inteligentes empleadas dentro y fuera del aula, deben ser implementados con moderación debido a las implicaciones éticas que conllevan (Bogina *et al*., 2022). La Unesco (2019), en el Consejo de Beijing sobre IA y educación, señala que los sectores educativos deben incorporar las directrices sobre IA en los marcos de competencias TIC, con el objetivo de respaldar la formación del personal docente en entornos altamente influenciados por la IA.
Luna Martínez (2023)	Es necesario reflexionar sobre la ética en un mundo donde la tecnología, como la IA y el teletrabajo, está cada vez más presente en nuestra vida diaria. Esta situación plantea el riesgo de que los humanos se vuelvan obsoletos y sean reemplazados por máquinas. La educación debe formar individuos que utilicen la tecnología de manera que promueva la responsabilidad social, la democracia, el respeto a la dignidad humana y la protección del medio ambiente, y contrarrestar las intenciones de control digital por parte del gran capital y gobiernos autoritarios.
Díaz-Guio *et al*. (2023)	Aunque las IA ya están disponibles en la educación en ciencias de la salud y, en menor medida, en la educación basada en simulación; aunque se esperan avances interesantes a mediano plazo, su implementación en la práctica cotidiana todavía enfrenta limitaciones económicas, contextuales, de interacción y éticas.
Mejía Azuero y Restrepo Ramírez (2023)	Es necesario aumentar y mejorar la investigación orientada a la toma de decisiones, particularmente en los aspectos pedagógicos y éticos relacionados con la transición hacia tecnologías digitales en la administración de justicia. Las investigaciones deben explorar la viabilidad y adecuación pedagógica del modelo virtual en la formación judicial, el impacto curricular en las competencias del juez actual, la coherencia entre el diseño formativo virtual y la satisfacción e impacto en los resultados institucionales, así como las implicaciones éticas y pedagógicas de integrar TIC e IA en la formación y práctica judicial.
Aguirre Sala (2023)	Los desafíos que plantea la IA pueden ser evaluados a través de distintos modelos y enfoques. Un enfoque relevante se centra en aspectos éticos, legales y culturales. Varias organizaciones y gobiernos han desarrollado códigos de ética para la IA, como el reciente código impulsado por el gobierno chino (Del Río, 2022) y las recomendaciones de la Unesco (2021). No obstante, estos códigos no son vinculantes, lo que significa que su adopción es opcional (Cortina, 2019; Lauer, 2021).
Gutiérrez-Cirlos *et al*. (2023)	Es urgente utilizar herramientas de detección de uso de IA (como GPTZero), así como definir políticas de uso de estas herramientas en universidades y hospitales, enseñar a los estudiantes a usar la IA de manera ética y responsable, e incluir materias de IA en los planes de estudio de las carreras de salud.
Morales Ramírez (2023)	Para reducir las amenazas asociadas con la IA, se deben realizar auditorías algorítmicas que incluyan detalles sobre los responsables del diseño, desarrollo e implementación del algoritmo, la metodología aplicada, los datos utilizados en el proceso de aprendizaje, las bases de datos empleadas y una definición clara de los grupos vulnerables que podrían ser afectados.
Ojeda *et al*. (2023)	ChatGPT se presenta como una herramienta útil para complementar y mejorar la enseñanza y el aprendizaje en el aula. Al enfrentar los desafíos y preocupaciones éticas vinculadas a su uso, los investigadores pueden utilizar el potencial de la IA de manera responsable para ampliar los horizontes del conocimiento humano y la comprensión (Ray, 2023).
Zielinski *et al*. (2023)	La recomendación 3 de la Asociación Mundial de Editores Médicos establece que los autores son responsables del contenido proporcionado por un *chatbot* en sus artículos, así como de asegurar su exactitud y la correcta atribución de todas las fuentes. Se deben atribuir, de forma adecuada, todas las fuentes citadas, incluso citas completas que apoyen las afirmaciones del *chatbot*. Además, los autores deben identificar el *chatbot* utilizado y la consulta específica realizada, así como explicar las medidas tomadas para mitigar el riesgo de plagio y asegurar una visión equilibrada y precisa en sus referencias.

Autor (año)	Resultados relativos
Palacios Gómez (2023)	Actualmente, los autores utilizan herramientas de escritura basadas en IA, como ChatGPT, Bard y Bing, sin reflexionar adecuadamente sobre sus limitaciones, lo que puede resultar en errores fácticos y de razonamiento en la escritura científica. Asimismo, los editores a menudo confían en los porcentajes de similitud proporcionados por algoritmos antiplagio como criterio para evaluar la originalidad de un manuscrito, lo que reemplaza el juicio de experto. Aunque la IA puede optimizar procesos, es la inteligencia humana la que define la creación, aplicación y el impacto de estas tecnologías en la sociedad.
Rodríguez Jiménez (2023)	Existen modalidades que se sitúan en el límite de las definiciones tradicionales de plagio académico, como la compra de trabajos a pedido que, aunque son trabajos originales, no son realizados por la persona que los presenta. La modalidad más reciente surge de la IA. ChatGPT es un generador de texto basado en IA capaz de producir, entre otras cosas, textos académicos originales. Estas modalidades se engloban bajo el concepto de fraude académico, una noción que abarca un mayor número de conductas indebidas.
De Vries (2023)	El uso de la IA para generar textos ha generado un debate sobre si los usuarios están incurriendo en plagio o en una nueva forma del mismo, denominada IAgio. Existe también la posibilidad de que los usuarios no plagien directamente, pero que el programa, al reutilizar fragmentos de textos de varios autores, termine plagiando contenidos disponibles en internet. La aparente solución es que los docentes pidan trabajos escritos a mano o utilicen programas que detecten el uso de IA (Truly, 2023), y verifiquen la originalidad del texto con herramientas como Turnitin (Rafter, 2023), además de asegurarse de que el trabajo presentado cumpla con los criterios de evaluación establecidos para la asignatura.
Cano *et al*. (2023)	En 2022, investigadores del Massachusetts General Hospital evaluaron a ChatGPT en el USMLE, un examen de licencia médica en EE.UU. ChatGPT respondió 350 preguntas y alcanzó el 60 % necesario para aprobar. Estos resultados fueron replicados en otros estudios, como en Yale y Dublín, donde el rendimiento varió según la dificultad de las preguntas, mostrando un menor desempeño en aquellas más difíciles. En otro experimento, Gao *et al*. recopilaron cincuenta resúmenes de cinco revistas médicas de alto impacto y pidieron a ChatGPT que generara resúmenes basándose en los títulos de investigación y las revistas correspondientes. Los resúmenes generados por IA presentaron menores índices de plagio en comparación con los escritos por investigadores. Tanto los científicos como los detectores de plagio e IA lograron identificar algunos de los resúmenes generados por ChatGPT, aunque ninguno de ellos fue completamente preciso.
Terrones Rodríguez y Rocha Bernardi (2024)	Las propuestas de gobernanza ética de la IA, fundamentadas en una perspectiva centrada en el ser humano, deben primero reconocer el valor educativo de la ética en la formación técnica. Además, deben incorporar herramientas intelectuales y habilidades de interacción basadas en la tradición ético-cívica dentro de los planes de estudio. La digitalización ha generado nuevas vulnerabilidades que exigen una mayor reflexividad y sensibilidad moral. Por ello, la promoción de la innovación educativa para crear espacios donde el conocimiento técnico y las humanidades se encuentren debe ser una parte integral de las políticas que buscan la gobernanza ética de la IA y el fortalecimiento de sociedades inclusivas, innovadoras y reflexivas en Europa.
Román-Acosta (2024)	A partir de una revisión de literatura, se identifican corrientes de pensamiento dominantes y áreas de controversia, destacándose además la complejidad del campo y la necesidad de investigaciones futuras que consideren tanto la filosofía como la ética en la evaluación del conocimiento producido por las herramientas tecnológicas.

Fuente: elaboración propia.

Inteligencia artificial y sesgos

La tabla 2 muestra las reflexiones que los autores formularon con relación a los sesgos incorporados en los algoritmos propios de la IA.

Tabla 2. Sesgos en la inteligencia artificial

Autor (año)	Hallazgos relativos
Roa Avella *et al*. (2022)	Los algoritmos predictivos que funcionan como una caja negra y contienen sesgos pueden comprometer los derechos y garantías procesales, especialmente cuando los tribunales priorizan la propiedad intelectual y el secreto empresarial. Organizaciones como el instituto AI Now trabajan para combatir estas injusticias algorítmicas y promover un diseño ético de algoritmos, destacando la necesidad de analizar su impacto en derechos y libertades, así como los sesgos y problemas de inclusión que afectan a diversas poblaciones. Los sesgos en los datos utilizados para diseñar y entrenar algoritmos pueden reflejar prejuicios sociales y perpetuar la discriminación contra minorías por etnia, género o estrato económico. Por lo tanto, es fundamental un diseño ético de los algoritmos que elimine estos sesgos y garantice la transparencia algorítmica, lo cual aseguraría que la IA pueda utilizarse eficazmente en la administración de justicia, con pleno respeto de los derechos y garantías procesales.
Calderón-Ortega y Cueto Calderón (2022)	Si se permitiera la utilización de la IA en los procesos judiciales, se comprometería una fase fundamental en la producción de pruebas: la valoración. En este contexto, se perdería la autenticidad, un criterio que se evalúa a través de las reglas de la lógica, la ciencia y la experiencia. Dicha prueba resultaría incompleta y podría introducir sesgos que afectarían la imparcialidad del decisor. El juez no podría incluirla en los fundamentos epistemológicos de su decisión. En caso de que el fallo se basara en esta prueba, la decisión estaría viciada de nulidad debido a la infracción del debido proceso probatorio, al haberse considerado un elemento jurídicamente inválido.
Gómez Rodríguez (2022)	La Declaración de Toronto sobre la protección del derecho a la igualdad y la no discriminación en los sistemas de aprendizaje automático (Amnistía Internacional, 2018) busca aplicar las normas internacionales de derechos humanos a estos sistemas. Señala que la IA, si se utiliza sin las salvaguardias adecuadas, tiene el potencial de exacerbar la discriminación y los sesgos preexistentes en ámbitos como el transporte, la atención sanitaria, la educación y la justicia. La Declaración hace un llamado a los Estados para que eviten la discriminación mediante la implementación de medidas de transparencia, responsabilidad y supervisión independiente en el diseño de la IA. Además, defiende el derecho a conocer el impacto de las neurotecnologías y a garantizar la integridad mental, comparándolo con el derecho a la protección frente a los sesgos de la IA.
Sánchez Acevedo (2022)	La gobernanza de la IA implica definir objetivos, regular su funcionamiento y verificar resultados, para enfrentar el *dilema de control* debido a la poca evidencia inicial. La Corporación Andina de Fomento propone una gobernanza responsable basada en anticipación, inclusión, adaptación y propósito, para equilibrar así la privacidad y la transparencia a fin de evitar sesgos y generar confianza. En Latinoamérica, algunos países han adoptado principios para garantizar equidad y transparencia en el uso de IA, previniendo discriminación. Es necesario que los sistemas de IA sean seguros, gestionados con un enfoque de riesgos y que permitan control externo, así como que consideren aspectos tecnológicos, legales y éticos. Los gobiernos utilizan IA para mejorar transparencia y rendición de cuentas, aunque su uso aún es incipiente y carece de aplicación clara de principios éticos. Las regulaciones deben ser claras, adaptarse a los avances tecnológicos y garantizar igualdad e inclusión para eliminar sesgos.
Gallent Torres *et al*. (2023)	Existen diversas preocupaciones relacionadas con el uso de la IA generativa y su impacto en la integridad académica, como el plagio y la creación de contenido no original; la utilización de herramientas para detectar textos generados por IA generativa; la implementación de métodos de evaluación alternativos; y la dependencia de estos sistemas, que podría debilitar el pensamiento crítico y dificultar la identificación de información incorrecta. Además, se advierte sobre la propagación de sesgos en los resultados generados. Un desafío ético adicional radica en la discriminación en la toma de decisiones automatizada, por ejemplo, en la selección de candidatos para programas académicos o en la evaluación del rendimiento estudiantil, donde los sesgos presentes en los datos de entrenamiento podrían influir. Por ello, es fundamental revisar los algoritmos para corregir estos sesgos y asegurar la equidad y la igualdad de oportunidades para todos.
García Peñalvo *et al*. (2023)	No es prudente suponer de manera ingenua que la tecnología es intrínsecamente neutral, ni que su impacto depende únicamente de quienes la crean y la utilizan. La tecnología no solo facilita la consecución de objetivos, sino que también interviene en la definición y conformación de esos mismos objetivos (Coeckelbergh, 2023). Por tanto, es necesario discutir sobre la toma de decisiones algorítmicas, su capacidad para influir y manipular (Flores-Vivar y García-Peñalvo, 2023a), los sesgos, las discriminaciones y las desigualdades resultantes (Holmes *et al*., 2022), la vigilancia, las competencias técnicas, las burbujas informativas y la exclusión social (Nemorin *et al*., 2023), así como la sustitución de seres humanos en los contextos del posthumanismo y el transhumanismo (Neubauer, 2021), y todas las interrelaciones entre estos aspectos. Estos temas son particularmente relevantes en el ámbito educativo, dado que inciden en el comportamiento humano. La IA tiene el potencial de influir tanto en la educación como en la sociedad en su conjunto, y se debe preparar a las personas para enfrentar un mundo cada vez más dominado por la tecnología, en el que la IA adquiere un papel preponderante (Flores-Vivar y García-Peñalvo, 2023b).

Autor (año)	Hallazgos relativos
Jiménez-García *et al*. (2023)	Es importante reconocer que la IA presenta limitaciones y desventajas que deben ser cuidadosamente consideradas. Aunque esta tecnología puede incrementar la eficiencia y la precisión en múltiples áreas, su uso inapropiado puede generar efectos adversos en la sociedad. En el ámbito educativo, resulta necesario abordar los desafíos éticos, morales y de privacidad que surgen en la gestión de datos personales, como lo señalan Renz y Vladova (2021) y Yuskovych-Zhukovska *et al*. (2022). Por ello, es necesario establecer políticas claras y medidas de protección que prevengan cualquier forma de discriminación o daño a los estudiantes. De esta manera, se pueden aprovechar los beneficios potenciales de la IA en la educación sin comprometer la privacidad ni la ética. Además, es imprescindible que tanto estudiantes como docentes desarrollen competencias adecuadas para la era de la IA, como destaca la Unesco (2021). Esto incluye la comprensión de los aspectos éticos y sociales, de la privacidad y la seguridad de los datos, así como de los impactos que esta tecnología puede tener en la sociedad (Arrieta *et al*., 2020).
Navarro-Dolmestch (2023)	El uso inadecuado de herramientas de IA puede conducir a un aprendizaje distorsionado, ya sea porque el estudiante adquiere información incorrecta, sesgada o discriminatoria; porque no logra internalizar y procesar adecuadamente dicha información; o porque las mejoras realizadas no reflejan su propio conocimiento en un texto corregido automáticamente (Perkins, 2023). La excesiva dependencia de herramientas de IA como fuente de información o para la realización de tareas, sin contrastarlas con otras fuentes ni realizar un análisis crítico, puede derivar en un aprendizaje profundamente distorsionado por los sesgos y la discriminación inherentes a la IA.
Luna Martínez (2023)	El uso sesgado de la tecnología con fines distintos al beneficio colectivo representa un riesgo para la humanidad, lo que nos lleva a cuestionarnos sobre el propósito y los límites de su uso. Es fundamental reflexionar críticamente sobre estos aspectos. En el contexto de la Cuarta Revolución Industrial, se debe considerar la ética en la implementación tecnológica, lo que podría llevar a la noción de una *ética tecnológica*. Proponer líneas formativas para una educación ética acorde a la actualidad es indispensable para enfrentar estos desafíos.
Mejía Azuero y Restrepo Ramírez (2023)	Aunque es necesaria una planificación estratégica para una transición hacia tecnologías digitales, algunos de los beneficios se intersectan con preocupaciones relacionadas con la pérdida de autonomía en la toma de decisiones por parte del juez. Más allá de las TIC tradicionales, es la IA la que plantea un desafío considerable para la objetividad judicial en la toma de decisiones. Esta inquietud es legítima, dado que un software como Compas, diseñado con complejas fórmulas algorítmicas para predecir la reincidencia delictiva, puede presentar fallos en cuanto a la transparencia e, incluso, incurrir en posibles sesgos étnicos y discriminación (Contini, 2020; Larson *et al*., 2016).
Aguirre Sala (2023)	Los riesgos y daños asociados con las decisiones de la IA se fundamentan en una base discriminatoria, derivada de datos sesgados que generan resultados inequitativos (Hacker, 2018). Existen pocas garantías de que los datos de entrada sean representativos y estén libres de prejuicios, lo que dificulta que los diseñadores aseguren que los resultados sean tanto éticos como legales. Un ejemplo de esto es la exclusión de mujeres en ciertas contrataciones laborales en Amazon (Dastin, 2018). La construcción y el etiquetado inadecuado de conjuntos de datos también pueden ser discriminatorios, lo que ha causado problemas como la invalidez jurídica de contratos inteligentes (Argelich, 2020) o errores graves en software de reconocimiento facial, como el incidente en el que Google Photos etiquetó a dos personas de color como gorilas (Zhang, 2015). Además, el sesgo histórico presente en los datos y en las representaciones intermediarias exacerba esta discriminación.
Gutiérrez-Cirlos *et al*. (2023)	ChatGPT presenta tanto ventajas como desventajas: aunque puede generar respuestas útiles, también puede incluir sesgos, referencias inexistentes y perpetuar estereotipos sexuales, como en búsquedas de matrimonio o delincuente en herramientas de búsqueda. Las preguntas sencillas, sin un entrenamiento previo, producen respuestas repetitivas y no siempre precisas, lo que hace necesaria la verificación humana. Por ello, se deben desarrollar reglas para asegurar la integridad, transparencia y originalidad de los textos enviados a revistas científicas.
Morales Ramírez (2023)	Los sesgos algorítmicos reflejan una sociedad dividida por prejuicios, que prioriza intereses distintos sobre las personas. Explicar la discriminación en sistemas de IA solo desde el punto de vista técnico es limitado, ya que la IA es un concepto sociotécnico influenciado por propósitos y relaciones sociales. Actualmente, no existen reglas para construir algoritmos libres de sesgos, pero las personas pueden identificar y corregir estos sesgos. Por ejemplo, un algoritmo que excluye a trabajadores del hogar no remunerados puede ser ajustado para incluirlos gracias a la intervención humana. Los algoritmos deben estar en constante revisión, con base en la dignidad y el respeto a la diversidad. Si se identifican fallas persistentes, los algoritmos deben ser desechados y reconstruidos. La diversidad y la auditoría algorítmica ayudarán a superar los prejuicios aceptados hasta ahora.

Autor (año)	Hallazgos relativos
Ojeda *et al*. (2023)	El uso de la IA en el contexto de la educación universitaria presenta varias limitaciones, entre las que se incluyen el potencial de sesgo y datos incompletos, la incapacidad para captar la complejidad de los sistemas biológicos, y las implicaciones éticas asociadas a estas tecnologías. Por lo tanto, es fundamental ser cautelosos frente a una dependencia excesiva de la IA en la medicina traslacional y garantizar que su uso sea responsable y ético (Van Dis *et al*., 2023). Además, es necesario establecer límites para estas herramientas, fortalecer las leyes existentes sobre discriminación y sesgo, y regular los usos peligrosos de la IA, lo cual contribuirá a asegurar un uso honesto, transparente y justo (Boutros *et al*., 2023).
Zielinski *et al*. (2023)	Los *chatbots* se activan mediante una instrucción en lenguaje natural, conocida como *prompt*, proporcionada por el usuario, y generan respuestas a través de modelos lingüísticos basados en estadísticas y probabilidades. Aunque estos resultados suelen ser lingüísticamente precisos y fluidos, presentan varias limitaciones. Los *chatbots* pueden incorporar sesgos, distorsiones, irrelevancias, tergiversaciones e incluso plagios, debido a los algoritmos que los gobiernan y a los materiales utilizados en su entrenamiento. Esto ha generado preocupación sobre los efectos de su uso en la creación y difusión del conocimiento, su capacidad para propagar y amplificar información errónea y desinformación, y su impacto en el empleo, la economía y la salud de las personas.
Ferrari (2023)	La universidad debe preservar su independencia respecto a industrias, empresas, organizaciones sociales y actores políticos. Es fundamental evitar la captura del investigador, una situación de dependencia que puede sesgar tanto las reflexiones como los resultados de la investigación, comprometiendo su calidad. Para lograrlo, es necesario que la universidad implemente políticas y normas claras que regulen de manera transparente su relación con estos sectores.
Palacios Gómez (2023)	Mejorar la capacidad de atención en los sistemas de salud, obtener diagnósticos más precisos, optimizar los tratamientos médicos y generar medidas de salud pública más eficientes son las promesas de los avances en IA. La Organización Mundial de la Salud reconoce estas expectativas, pero evidencia también la necesidad de garantizar la transparencia, la explicación y la comprensión de cada aplicación de IA en el ámbito de la salud, con una evaluación constante que asegure la equidad, la inclusión y la sostenibilidad.
Cano *et al*. (2023)	Al revisar las normas de autoría del Comité Internacional de Editores de Revistas Médicas, se establece que es considerado autor de un manuscrito quien contribuye a su redacción. En este contexto, se ha discutido si la IA podría ser incluida como coautora, pero esta idea ha sido rechazada, dado que los algoritmos no pueden asumir la responsabilidad por lo escrito. No obstante, es obligatorio declarar el uso de IA al momento de publicar el manuscrito. Se exige que cada sección redactada con apoyo de IA sea revisada críticamente por el equipo de investigación para verificar su precisión, sesgo, relevancia y solidez en el razonamiento. La IA es una herramienta novedosa que no sustituye el papel del investigador ni el pensamiento científico, y su uso debe ser debidamente declarado al momento de la publicación.
Terrones Rodríguez y Rocha Bernardi (2024)	Las iniciativas impulsadas tanto por la Comisión Europea como por el Gobierno de España reflejan un compromiso con la gobernanza ética de la IA. Sin embargo, a pesar de que estos enfoques son valiosos para establecer un ecosistema basado en la confianza, carecen de una estrategia educativa específica dentro del ámbito de la educación superior. En particular, los programas de estudios en ingeniería no integran de manera suficiente componentes formativos en ética aplicada. Esta deficiencia obstaculiza la comprensión de los efectos controvertidos y ambivalentes de los sistemas sociotécnicos, y limita el desarrollo del juicio crítico y la sensibilidad moral necesarios para reconocer aspectos que podrían ser ignorados debido al predominio del pensamiento computacional y al aceleracionismo económico.
Román-Acosta (2024)	La precisión y la validez del conocimiento producido por máquinas dependen de factores como la transparencia, la confiabilidad y las consideraciones éticas. A medida que la IA continúa desarrollándose, se torna imprescindible crear marcos de evaluación que garanticen no solo la exactitud del conocimiento generado, sino también su solidez ética. Para enfrentar los retos que plantea este panorama en transformación, es necesario un enfoque interdisciplinario que involucre la colaboración entre epistemólogos, éticos y tecnólogos.

Fuente: elaboración propia.

Resultado de la experimentación con IA

El experimento consistió en utilizar ChatGPT4o para redactar resúmenes críticos de sentencias dictadas por el Tribunal Constitucional del Perú. Cada resumen debía contener: una síntesis de los hechos relevantes, la identificación de la controversia, la base normativa pertinente, la razón esencial de la decisión y la parte resolutiva.

ChatGPT4o generó los resúmenes solicitados siguiendo las directrices proporcionadas. Luego, los textos resultantes fueron sometidos a Turnitin, una herramienta de detección de similitud. El resultado mostró un 27 % de plagio, concentrada principalmente en las citas de las normas jurídicas. Este nivel puede considerarse aceptable, dado que las citas legales suelen ser fijas y recurrentes.

Los textos también se analizaron con ChatGPTZero (no asociada a OpenAI), una herramienta diseñada para detectar la autoría humana o generada por IA. Dicho sistema determinó que no podía dictaminar con certeza la autoría de los textos. Sin embargo, consideró que la mayor probabilidad era que los textos hubieran sido redactados por humanos.

Adicionalmente, para evaluar la capacidad de análisis crítico de ChatGPT4o, se le requirió analizar una investigación sobre trabajo infantil (Miranda Huayamares, 2018). Primero se le brindó la instrucción de analizar el texto con base en una rúbrica, con el siguiente comando: "Necesito que actúes con el rigor académico de un par evaluador de alto nivel. Aplica la rúbrica para evaluar este trabajo y así identificar todas las debilidades que puedas encontrar". Los comentarios de mejora que brindó el sistema fueron: "La portada necesita incluir un título abreviado y una declaración de conflicto de interés"; "El resumen debe ser conciso y se debe agregar la traducción al inglés"; "La introducción debe mencionar explícitamente los objetivos del estudio"; "La metodología debe incluir referencias a los estándares éticos"; "Las tablas y citas deben ajustarse al formato APA".

Sin embargo, en un segundo momento se le brindó esta instrucción: "Ahora céntrate en el contenido. ¿Se pueden extraer conclusiones generales sobre un tema tan sensible como el abordado a partir de cuatro casos aleatoriamente seleccionados y cuya información ha sido tomada de fuentes no arbitradas? S*é* detallado y profundo en el análisis y la discusión debe contrastar más las posturas diversas". El resultado, organizado en subtítulos y viñetas, arrojó un dictamen de ChatGPTZero de mayor probabilidad de haber sido redactado por un humano. Luego se le añadió la instrucción de parafrasear el mismo contenido, pero de forma más narrativa sin subtítulos ni viñetas, evitando fórmulas repetitivas e incluyendo alguna inconsistencia gramatical menor. El resultado fue el siguiente:

> El trabajo de Claudia Cecilia Miranda Huayamares aborda un tema tan sensible y complejo como los aspectos positivos del trabajo infantil. A continuación, se ofrece un análisis detallado sobre si es posible extraer conclusiones generales a partir de los casos presentados y las fuentes utilizadas.
>
> Los casos de Michael Sayman, Moshe Kai Cavalin, Jaequan Faulkner y José Quisocala son ejemplos excepcionales de niños que han tenido éxito en sus respectivos campos desde una edad temprana. Michael Sayman, por ejemplo, comenzó a desarrollar aplicaciones móviles exitosas desde los 13 años, y Moshe Kai Cavalin ingresó a la universidad a los 8 años y ya trabaja para la NASA. Jaequan Faulkner inició un puesto de hot dogs a los 13 años y José Quisocala creó un banco basado en residuos reciclados.
>
> Sin embargo, las fuentes de estos casos provienen de medios de comunicación general y no de revistas académicas revisadas por pares. Esto limita la confiabilidad y validez de la información, ya que no han pasado por un proceso riguroso de revisión académica. Los artículos periodísticos pueden tener sesgos y no siempre presentan una visión completa o crítica del tema.
>
> Además, los casos seleccionados parecen ser anecdóticos y excepcionales, representando situaciones atípicas de éxito extremo que no reflejan la realidad de la mayoría de los niños que trabajan, quienes enfrentan condiciones adversas y riesgos significativos para su salud y educación. No hay evidencia de un proceso sistemático o criterios rigurosos para la selección de estos casos, lo que puede introducir un sesgo significativo.
>
> Generalizar los resultados de estos casos a la población general de niños trabajadores es metodológicamente incorrecto. Las conclusiones generales sobre los beneficios del trabajo infantil basadas en estos casos pueden ser engañosas y no considerar el daño potencial y los riesgos involucrados. Para abordar un tema tan sensible y complejo, se necesitarían estudios empíricos basados en muestras

representativas, datos longitudinales y revisiones de literatura académica arbitrada.

El trabajo menciona el contexto cultural y económico, pero no profundiza en cómo estos factores influyen en la percepción y realidad del trabajo infantil en diferentes regiones. Es vital considerar cómo las condiciones socioeconómicas y culturales específicas afectan la vida de los niños trabajadores y la aplicabilidad de las conclusiones en diferentes contextos.

En resumen, el trabajo presenta una visión sesgada y limitada al basarse en casos anecdóticos y fuentes no arbitradas. Las conclusiones generales sobre los aspectos positivos del trabajo infantil no son válidas debido a la falta de rigor metodológico y la no representatividad de los casos seleccionados. Para hacer afirmaciones más fundamentadas y precisas, es necesario utilizar fuentes académicas revisadas por pares, una muestra representativa de casos y un análisis más profundo y crítico del contexto en el que se desarrolla el trabajo infantil.

El dictamen de ChatGPTZero se mantuvo invariable, con mayor probabilidad de haber sido redactado por un ser humano. Por su parte, el reporte generado por Turnitin arrojó un 0 % de similitud.

Iniciativas institucionales

Un elemento clave en esta materia es la capacitación ética y técnica de estudiantes y docentes. Así, Unesco (2020) publicó el primer borrador de un marco de competencias en IA, con directrices específicas para estudiantes y profesores, abarcando fundamentos técnicos, ética y gobernanza.

Igualmente, la Comisión Europea (2022) elaboró el documento *Directrices* éticas sobre el uso de la inteligencia artificial y los datos en la educación, orientado a asegurar la integridad académica y proteger los derechos fundamentales de estudiantes y docentes. Dichas directrices resaltan aspectos como la transparencia en el funcionamiento de los sistemas de IA, la equidad e inclusión para estudiantes con necesidades especiales, la importancia de la supervisión humana para prevenir dependencias excesivas, así como la privacidad y la gobernanza rigurosa de los datos personales.

Además, según una encuesta realizada por la American Bar Association (2025), el 55 % de las facultades de derecho encuestadas ofrecen cursos específicamente diseñados para enseñar IA, mientras que un 83 % ofrece oportunidades curriculares prácticas, como clínicas jurídicas, para aprender a utilizar herramientas de IA en contextos reales. Además, un 69 % de estas instituciones ha actualizado sus políticas de integridad académica en respuesta a la IA generativa, y un 62 % ya integra formalmente contenidos sobre IA en el primer año del currículo académico. El 93 % de las facultades planea cambios adicionales en sus programas educativos debido a la creciente relevancia de la IA.

En el ámbito legal, la American Bar Association (2024) emitió la Formal Opinion 512, que establece orientaciones detalladas sobre el uso ético y responsable de herramientas de IA generativa en la práctica jurídica. Esta opinión resalta que los abogados deben adquirir y mantener una competencia técnica razonable sobre las capacidades y limitaciones específicas de las herramientas de IA que utilizan, así como realizar una evaluación continua debido a su rápido desarrollo. Además, hace hincapié en la obligación de proteger estrictamente la confidencialidad de la información de los clientes, por lo que se debe exigir consentimiento informado antes de ingresar datos sensibles en herramientas de IA. También desarrolla la importancia de una comunicación transparente con los clientes acerca del uso de estas tecnologías, particularmente cuando las decisiones derivadas de la IA puedan afectar el curso de la representación. Finalmente, señala la necesidad de supervisar adecuadamente a todo el personal que use herramientas de IA, asegurar la veracidad de la información proporcionada por dichas herramientas ante tribunales y mantener criterios razonables al fijar honorarios vinculados al uso de estas tecnologías.

Discusión

La tabla 1 revela importantes aspectos sobre la relación entre la IA y la integridad académica. Los trabajos analizados destacan tanto los beneficios potenciales como los riesgos asociados con su uso en la educación jurídica. Dichos resultados

demuestran la importancia de la ética y la responsabilidad en la implementación de estas tecnologías, lo que evidencia la necesidad de un enfoque crítico y reflexivo por parte de educadores y estudiantes.

Uno de los aspectos más destacados es la preocupación por el uso indebido de herramientas de IA, como el plagio y la suplantación de autoría. Los estudios señalan que la disponibilidad y facilidad de uso de estas tecnologías pueden tentar a los estudiantes a emplearlas de manera deshonesta, lo que compromete los valores fundamentales de la educación. Navarro-Dolmestch (2023) y Gallent Torres *et al*. (2023) discuten cómo el fraude académico puede proliferar si no se implementan controles adecuados y se educa a los estudiantes sobre la importancia de la honestidad académica.

La integridad académica también se ve amenazada por la falta de políticas claras y medidas de supervisión efectivas. García Peñalvo *et al*. (2023) destacan la necesidad de una coordinación internacional para maximizar los beneficios de la IA mientras se minimizan los riesgos de su uso fraudulento. La promoción de buenas prácticas y la detección del fraude académico son esenciales para mantener la igualdad y la formación adecuada de los estudiantes.

Por otro lado, también se han abordado aspectos positivos de la IA en la educación, como su capacidad para mejorar la enseñanza y el aprendizaje mediante la personalización y la retroalimentación automatizada. Sin embargo, Jiménez-García *et al*. (2023) advierten que se deben abordar los desafíos éticos y de privacidad asociados con el manejo de datos personales. La implementación de políticas claras y medidas de protección puede permitir a las instituciones educativas aprovechar los beneficios de la IA sin comprometer la privacidad y la ética.

La importancia de un enfoque equilibrado también es abordada por Norman-Acevedo (2023), quien sostiene que, aunque la IA puede ser una herramienta valiosa para mejorar el aprendizaje, es fundamental considerar tanto sus beneficios como sus desventajas. Solo se ha comenzado a explorar el potencial de la IA, y es esencial actuar con precaución para no quedarse atrás en su implementación, pero siempre manteniendo un equilibrio entre la innovación tecnológica y la preservación de los valores académicos.

Ahora, con relación a los sesgos en la IA, y conforme a los resultados plasmados en la tabla 2, es posible identificar varios riesgos implícitos, así como estrategias potenciales para abordarlos. Uno de tales riesgos tiene que ver con la discriminación y la desigualdad. Los algoritmos predictivos pueden perpetuar y amplificar los sesgos presentes en los datos de entrenamiento, con resultados discriminatorios contra minorías étnicas, de género o socioeconómicas. Roa Avella *et al*. (2022) señalan que estos sesgos pueden comprometer los derechos y garantías procesales, con el riesgo de perpetuar injusticias en el sistema judicial.

También resulta problemática la falta de transparencia del algoritmo. Calderón-Ortega y Cueto Calderón (2022) señalan que la opacidad de los algoritmos, a menudo considerados como cajas negras, puede impedir la comprensión y evaluación adecuada de las decisiones tomadas por IA. Esto puede llevar a decisiones judiciales basadas en información no verificable, comprometiendo la legitimidad del proceso judicial. Por otro lado, Gómez Rodríguez (2022) se enfoca en la reproducción de prejuicios sociales y advierte que la IA, sin salvaguardas adecuadas, puede amplificar los prejuicios existentes, lo que afecta a sectores clave como la justicia, la educación y la atención médica. En su estudio aborda a la Declaración de Toronto, que destaca la necesidad de prevenir la discriminación mediante la transparencia y la supervisión independiente. A su vez, la falta de claridad en la gobernanza de la IA puede generar desconfianza entre los usuarios y el público en general. A su vez, Sánchez Acevedo (2022) afirma que la poca evidencia inicial y la implementación sin regulaciones claras pueden socavar la confianza en los sistemas de IA.

Frente a dichos riesgos se han podido identificar estrategias para mitigarlos. Así, con relación al diseño ético y transparente de algoritmos, se requiere implementar principios éticos en el diseño de algoritmos para garantizar la transparencia y la equidad. Roa Avella *et al*. (2022) recomiendan un diseño ético de los algoritmos que elimine los sesgos y asegure la transparencia algorítmica. A su vez, Gómez Rodríguez (2022) sugiere que los

Estados deben establecer mecanismos de supervisión independiente y rendición de cuentas para garantizar que los sistemas de IA operen de manera justa y equitativa. Esto incluye la adopción de normas que obliguen a los desarrolladores a revelar el funcionamiento interno de sus algoritmos. Complementariamente, Sánchez Acevedo (2022) propone una gobernanza de la IA basada en la anticipación, inclusión, adaptación y propósito, que equilibre privacidad y transparencia para evitar sesgos. La implementación de regulaciones claras que se adapten a los avances tecnológicos es fundamental para garantizar la equidad y la inclusión.

También se necesita educar tanto a desarrolladores como a usuarios sobre los riesgos de los sesgos en la IA. Calderón-Ortega y Cueto Calderón (2022) explican la importancia de capacitar a los jueces y otros profesionales del derecho para que comprendan las limitaciones y posibles sesgos de las pruebas generadas por IA.

Otra estrategia de mitigación que resalta es la de auditorías algorítmicas. Morales Ramírez (2023) destaca la necesidad de auditorías algorítmicas continuas para identificar y corregir sesgos. Estas deben incluir información detallada sobre el diseño, desarrollo e implementación de los algoritmos, para asegurar que se respeten la dignidad y la diversidad.

Así, la mitigación de los riesgos asociados con los sesgos en la IA requiere un enfoque multifacético que incluya el diseño ético y transparente de algoritmos, la supervisión independiente, la gobernanza responsable, la capacitación y la realización de auditorías algorítmicas continuas. Estas estrategias contribuirán a que la IA se utilice de manera justa y equitativa, con pleno respeto de los derechos y garantías procesales.

Ahora, con relación al experimento exploratorio, el resultado de 27 % de similitud en Turnitin, concentrado principalmente en las citas de normas, sugiere que la IA es capaz de generar contenido relevante con un nivel aceptable de originalidad. Esto incide en el ámbito académico y profesional, donde la integridad y la originalidad son pilares fundamentales. La capacidad de la IA para producir textos que cumplen con estándares aceptables de similitud muestra su potencial para asistir en tareas de redacción y análisis, así como para retar a los sistemas educativos tradicionales.

Por su parte, el dictamen de ChatGPTZero, que indica una alta probabilidad de que el texto sea redactado por humanos, destaca otro aspecto importante: la fluidez y coherencia del texto generado por la IA. Esto sugiere que la IA no solo puede producir contenido original, sino también hacerlo de una manera que imite eficazmente el estilo y la estructura de la redacción humana. Esta capacidad tiene implicaciones significativas para la educación jurídica, pues la claridad y precisión en la redacción son características importantes, al tiempo que demanda una redefinición del rol evaluador del docente.

Como se aprecia, el uso de IA en la redacción de textos jurídicos plantea varias cuestiones éticas y educativas. Por un lado, la capacidad de la IA para generar textos de alta calidad puede mejorar la eficiencia y la productividad. Por otro lado, existe el riesgo de que los estudiantes dependan excesivamente de estas herramientas, lo que podría afectar su desarrollo de habilidades críticas y analíticas. La integración de la IA en la educación debe, por lo tanto, ir acompañada de una educación adecuada sobre su uso ético y responsable.

Este estudio demuestra la necesidad de equilibrar los beneficios de la IA con los riesgos asociados. El experimento demuestra que, aunque la IA puede ser una herramienta poderosa para la redacción y el análisis, se necesita de la implementación de políticas claras y medidas de protección para garantizar la integridad académica. Esto incluye educar a los estudiantes sobre la importancia de la honestidad académica y desarrollar mecanismos para detectar y prevenir el uso indebido de la IA.

El experimento presentado pone de relieve tanto el potencial como los desafíos de integrar la IA en la formación jurídica. Si se utiliza de manera ética y responsable, la IA puede enriquecer la educación jurídica, mejorar la eficiencia y apoyar el desarrollo de habilidades críticas. Sin embargo, se debe mantener un enfoque equilibrado que promueva tanto la innovación tecnológica como la preservación de los valores académicos fundamentales. Se requiere de posteriores estudios interdisciplinarios

en esta materia orientados a diseñar estrategias eficaces para hacer frente a estos retos.

El análisis del experimento adicional realizado con ChatGPT-4o, donde se evaluó su capacidad para analizar críticamente una investigación sobre trabajo infantil, ofrece resultados relevantes y plantea importantes reflexiones sobre las herramientas actuales de detección de plagio y autoría. En este experimento, ChatGPT-4o generó un análisis detallado y coherente de la investigación sobre trabajo infantil, y siguió instrucciones específicas para evaluar la metodología y la validez de las conclusiones del estudio. Los resultados fueron evaluados mediante dos herramientas: Turnitin y ChatGPTZero. Turnitin mostró un 0 % de similitud, lo que indica una ausencia total de plagio, mientras que ChatGPTZero determinó que la respuesta tenía una alta probabilidad de haber sido redactada por un humano.

El resultado obtenido de Turnitin, con un 0 % de similitud, sugiere que esta herramienta podría estar volviéndose obsoleta en la detección de contenido generado por IA avanzada. Turnitin, diseñado para identificar plagio mediante la comparación con una base de datos de textos previamente registrados, puede no ser suficiente para detectar textos originales generados por IA que no copian contenido existente pero que, no obstante, podrían no reflejar una autoría genuina.

Nos enfrentamos así a un escenario en el que la redacción por IA es indetectable, por lo que las instituciones educativas deben adoptar enfoques estratégicos y multifacéticos para abordar los desafíos de integridad académica. Se debe entonces rediseñar los métodos de evaluación, incrementar el uso de presentaciones orales y evaluaciones presenciales, en los que los estudiantes expliquen y defiendan sus trabajos en persona. También es útil realizar más evaluaciones y trabajos escritos en clase, donde se pueda supervisar directamente el proceso de redacción, y fomentar proyectos colaborativos que requieran interacciones regulares y contribuciones continuas de los estudiantes.

El desarrollo de competencias críticas y creativas es fundamental. Esto incluye la integración de actividades y ejercicios que promuevan el pensamiento crítico y la capacidad de análisis profundo, habilidades difíciles de replicar por la IA. Además, se debe valorar y calificar el proceso de investigación y redacción, incluyendo borradores y versiones intermedias, no solo el producto final.

El fortalecimiento de la honestidad académica necesita también de la revisión y actualización de los códigos de ética académica, con inclusión de directrices claras sobre el uso de IA. Al mismo tiempo, se deben implementar programas de educación ética que aborden el uso responsable de la tecnología y sus implicaciones sobre la integridad académica. Las instituciones deben invertir en el desarrollo y adopción de herramientas avanzadas que detecten no solo plagio, sino también patrones de redacción característicos de IA, y realizar auditorías periódicas de trabajos académicos para identificar tendencias y posibles casos de uso indebido.

También es indispensable fomentar la autenticidad en el aprendizaje. Puede resultar muy beneficioso diseñar proyectos y tareas que requieran la incorporación de experiencias personales, perspectivas únicas y conocimientos específicos del contexto local del estudiante, así como incluir componentes reflexivos donde los estudiantes expliquen sus procesos de aprendizaje y la manera como llegaron a sus conclusiones.

A la par, es necesario adaptar el currículo para integrar la IA en la enseñanza, así como ayudar a los estudiantes a comprender su funcionamiento y limitaciones, y desarrollar competencias digitales avanzadas que los preparen para el uso ético y efectivo de la IA en sus campos profesionales.

En el ámbito institucional, iniciativas como las de la Unesco (2020) y la Comisión Europea (2022) han establecido marcos éticos y técnicos claros para la integración de la IA en la educación, con énfasis en la importancia de la transparencia, la equidad y la inclusión. La American Bar Association (2024) complementa estos esfuerzos con orientaciones específicas sobre el uso responsable de IA generativa en la práctica legal, y que resaltan la necesidad de mantener la competencia técnica, proteger la confidencialidad, obtener consentimiento informado de los clientes y asegurar la transparencia en su uso.

Estos aportes reflejan un consenso creciente sobre la necesidad de integrar la IA en la educación

jurídica con cautela, ética y responsabilidad, considerando tanto su potencial innovador como sus posibles impactos negativos. En definitiva, el reto para las instituciones educativas y profesionales es implementar estas tecnologías de manera equilibrada, aprovechando sus ventajas educativas mientras se minimizan sus riesgos éticos y sociales.

Conclusiones

La IA puede optimizar diversos aspectos de la educación jurídica, que incluyen la alfabetización digital y las competencias tecnológicas de estudiantes y docentes. La personalización del aprendizaje y la retroalimentación automatizada permiten una enseñanza más adaptada y efectiva.

Sin embargo, la IA también presenta riesgos para la integridad académica, como el plagio y la suplantación de autoría. Se deben implementar controles adecuados y educar a los estudiantes sobre la importancia de la honestidad académica para mitigar estos riesgos.

La incorporación de la IA en la educación jurídica requiere el desarrollo de marcos normativos y éticos idóneos y proporcionales. Es indispensable contar con regulaciones claras y medidas de supervisión efectivas, para proteger los derechos de los individuos y asegurar un uso responsable de estas tecnologías.

La IA puede perpetuar y amplificar sesgos presentes en los datos de entrenamiento, con resultados marcados por la discriminación y desigualdad. Se deben diseñar algoritmos éticos y transparentes, realizar auditorías continuas y capacitar a los desarrolladores y usuarios para identificar y corregir estos sesgos.

La integración de la IA debe equilibrar la innovación tecnológica con la preservación de valores académicos fundamentales. Es necesario actuar con precaución y mantener un enfoque crítico para evitar quedarse atrás en su implementación sin comprometer la ética y la responsabilidad.

La capacidad de la IA para automatizar tareas administrativas y proporcionar retroalimentación en tiempo real sugiere la necesidad de redefinir el rol del docente, enfocándose en actividades más estratégicas y pedagógicas. Así, al contrario de lo que se podría pensar a partir de estas nuevas tecnologías, el rol docente no queda relativizado, sino más bien reafirmado.

Los resultados experimentales indican que la IA puede generar contenido de calidad con un nivel aceptable de originalidad y coherencia. Sin embargo, es esencial continuar con investigaciones futuras para diseñar estrategias efectivas que aborden los desafíos éticos y educativos asociados con su uso.

Por último, asumir que la redacción por IA es indetectable implica un cambio significativo en la forma como las instituciones educativas abordan la enseñanza y la evaluación. No se trata solo de detectar el uso de IA, sino de promover un ambiente académico donde la autenticidad, la ética y el aprendizaje profundo sean valorados y fomentados, con una mezcla de enfoques pedagógicos, tecnológicos y éticos para asegurar que los estudiantes desarrollen habilidades genuinas y se adhieran a los principios de integridad académica.

Referencias

Aguirre Sala, J. F. (2023). Modelos y buenas prácticas evaluativas para detectar impactos, riesgos y daños de la inteligencia artificial. *Paakat: Revista de Tecnología y Sociedad*, *12*(23), 1–20. https://doi.org/10.32870/Pk.a12n23.742

American Bar Association. (2024). Formal Opinion 512. Generative Artificial Intelligence Tools. https://www.americanbar.org/content/dam/aba/administrative/professional_responsibility/ethics-opinions/aba-formal-opinion-512.pdf

American Bar Association Survey. (2025). AI and Legal Education Survey Results 2024. https://www.americanbar.org/content/dam/aba/administrative/office_president/task-force-on-law-and-artificial-intelligence/2024-ai-legal-ed-survey.pdf

Ayuso del Puerto, D., y Gutiérrez Esteban, P. (2022). La inteligencia artificial como recurso educativo durante la formación inicial del profesorado. *RIED-Revista Iberoamericana de Educación a Distancia*, *25*(2), 347–362. https://doi.org/10.5944/ried.25.2.32332

Calderón-Ortega, M. A., y Cueto Calderón, C. A. (2022). Prueba por inteligencia artificial: una propuesta de producción probatoria desde el dictamen pericial científico en Colombia. *Civilizar: Ciencias Sociales y Humanas*, *22*(42). https://doi.org/10.22518/jour.ccsh/20220106

Calvo, P. (2022). Una ética de la investigación en el marco de las éticas aplicadas. *Veritas*, (52), 29–51. https://www.scielo.cl/pdf/veritas/n52/0718-9273-veritas-52-29.pdf

Cano, F., Schonhaut, B. L., Harris, P. R., Millán, K. T., Mena, N. P., Lizama, C. M., … Weisstaub, G. (2023). Andes Pediatrica y publicaciones científicas en la era de la inteligencia artificial. *Andes Pediatrica*, *94*(6), 667–671. https://doi.org/10.32641/andespediatr.v94i6.4974

Comisión Europea. (2022). *Directrices* éticas *sobre el uso de la inteligencia artificial y los datos en la educación y formación para los educadores*. https://op.europa.eu/en/publication-detail/-/publication/d81a0d54-5348-11ed-92ed-01aa75ed71a1/language-en

De Vries, W. (2023). Como (no) combatir el fraude académico. Lecciones internacionales. *Revista Mexicana de Investigación Educativa*, *28*(97), 637–650. https://www.scielo.org.mx/pdf/rmie/v28n97/1405-6666-rmie-28-97-637.pdf

Díaz-Guio, D. A., Henao, J., Pantoja, A., Arango, M. A., Díaz-Gómez, A. S., y Camps Gómez, A. (2023). Inteligencia artificial, aplicaciones y desafíos en la educación basada en simulación. *Colombian Journal of Anesthesiology*, (52), 1–6. https://doi.org/10.5554/22562087.e1085

Ferrari, C. A. (2023). Por una universidad para una nueva sociedad y economía. *Revista de Economía Institucional*, *25*(48), 143–176. https://doi.org/10.18601/01245996.v25n48.09

Forero-Corba, W., y Negre Bennasar, F. (2023). Técnicas y aplicaciones de machine learning e inteligencia artificial en educación: una revisión sistemática. *RIED-Revista Iberoamericana de Educación a Distancia*, *27*(1), 209–253. https://doi.org/10.5944/ried.27.1.37491

Gallent Torres, C., Zapata González, A., y Ortego Hernando, J. L. (2023). El impacto de la inteligencia artificial generativa en educación superior: una mirada desde la ética y la integridad académica. *RELIEVE - Revista Electrónica de Investigación y Evaluación Educativa*, *29*(2). https://doi.org/10.30827/relieve.v29i2.29134

Galvis Martínez, J. C., y Esquivel Zambrano, L. M. (2022). Derechos y deberes en la inteligencia artificial: dos debates inconclusos entorno a su regulación. *Nuevo Derecho*, *18*(31), 1–17. https://doi.org/10.25057/2500672X.1479

García Peñalvo, F. J., Llorens-Largo, F., y Vidal, J. (2023). La nueva realidad de la educación ante los avances de la inteligencia artificial generativa. *RIED-Revista Iberoamericana de Educación a Distancia*, *27*(1), 9–39. https://doi.org/10.5944/ried.27.1.37716

Gómez Rodríguez, J. M. (2022). Inteligencia artificial y neuroderechos. Retos y perspectivas. *Cuestiones Constitucionales*, (46), 93–119. https://revistas.juridicas.unam.mx/index.php/cuestiones-constitucionales/article/view/17049/17593

Gutiérrez-Cirlos, C., Carrillo-Pérez, D. L., Bermúdez-González, J. L., Hidrogo-Montemayor, I., Carrillo-Esper, R., y Sánchez-Mendiola, M. (2023). ChatGPT: oportunidades y riesgos en la asistencia, docencia e investigación médica. *Gaceta Médica de México*, (159), 382–389. https://doi.org/10.24875/GMM.230001671

Jiménez-García, E., Orenes-Martínez, N., y López-Fraile, L. A. (2023). Rueda de la pedagogía para la inteligencia artificial: adaptación de la rueda de Carrington. *RIED-Revista Iberoamericana de Educación a Distancia*, *27*(1), 87–113. https://doi.org/10.5944/ried.27.1.37622

Luna Martínez, A. (2023). Ética docente frente a la revolución tecnológica (CRI). Una perspectiva hermenéutica-analógica. *Revista CoPaLa. Construyendo Paz Latinoamericana*, *8*(18), 1–14. https://www.redalyc.org/articulo.oa?id=668174966010

Mata García, B., Santos Cervantes, C., y Zepeda Moreno, M. E. (2024). Sociedades automatizadas y educación 4.0. Retos, perspectivas y contradicciones de pensar la formación humana como ingeniería social. *Revista Latinoamericana de Estudios Educativos*, *54*(1), 165–188. https://doi.org/10.48102/rlee.2024.54.1.613

Mejía Azuero, J. C., y Restrepo Ramírez, A. (2023). Transición digital en la formación y práctica judicial: beneficios y desafíos. *Revista de Derecho*, (60), 90–108. https://doi.org/10.14482/dere.60.001.416

Miranda Huayamares, C. C. (2018). *Trabajo infantil: Aspectos positivos de las labores realizadas por los niños* [Tesis de segunda especialidad, Pontificia Universidad Católica del Perú]. https://tesis.pucp.edu.pe/repositorio/handle/20.500.12404/13719

Morales Ramírez, G. (2023). Problemática antropológica detrás de la discriminación generada a partir de los algoritmos de la inteligencia artificial. *Medicina y Ética*, *34*(2), 429–455. https://doi.org/10.36105/mye.2023v34n2.04

Navarro-Dolmestch, R. (2023). Descripción de los riesgos y desafíos para la integridad académica de aplicaciones generativas de inteligencia artificial. *Derecho PUCP*, (91), 231–270. https://doi.org/10.18800/derechopucp.202302.007

Norman-Acevedo, E. (2023). La inteligencia artificial en la educación: una herramienta valiosa para los tutores virtuales universitarios y profesores universitarios. *Panorama*, *17*(32), 1–11. https://doi.org/10.15765/pnrm.v17i32.3681

Ojeda, A. D., Solano-Barliza, A. D., Ortega Álvarez, D., y Boom Cárcamo, E. (2023). Análisis del impacto de la inteligencia artificial ChatGPT en los procesos de enseñanza y aprendizaje en la educación universitaria. *Formación Universitaria*, *16*(6), 61–70. https://doi.org/10.4067/S0718-50062023000600061

Palacios Gómez, M. (2023). Inteligencia humana para autores, revisores y editores que utilicen inteligencia artificial. *Colombia Medica*, *54*(3), e1005867. https://doi.org/10.25100/cm.v54i3.5867

Roa Avella, M. del P., Sanabria-Moyano, J. E., y Dinas-Hurtado, K. (2022). Uso del algoritmo Compas en el proceso penal y los riesgos a los derechos humanos. *Revista Brasileira de Direito Processual Penal*, *8*(1), 275–310. https://doi.org/10.22197/rbdpp.v8i1.615

Rodríguez Jiménez, J. R. (2023). Ampliando el horizonte sobre el plagio académico. *Revista Mexicana de Investigación Educativa*, *28*(97), 661–672. https://www.scielo.org.mx/pdf/rmie/v28n97/1405-6666-rmie-28-97-661.pdf

Román-Acosta, D. (2024). Exploración filosófica de la epistemología de la inteligencia artificial: Una revisión sistemática. *Uniandes Episteme*, *11*(1), 101–122. https://doi.org/10.61154/rue.v11i1.3388

Sánchez Acevedo, M. E. (2022). La inteligencia artificial en el sector público y su límite respecto de los derechos fundamentales. *Estudios Constitucionales*, *20*(2), 257–284. https://doi.org/10.4067/S0718-52002022000200257

Terrones Rodríguez, A. L., y Rocha Bernardi, M. (2024). El valor de la ética aplicada en los estudios de ingeniería en un horizonte de inteligencia artificial confiable. *Sophía*, (36), 221–245. https://doi.org/10.17163/soph.n36.2024.07

Unesco. (2020). *First Draft of the Recommendation on the Ethics of Artificial Intelligence.* https://unesdoc.unesco.org/ark:/48223/pf0000373434

Zielinski, C., Winker, M. A., Aggarwal, R., Ferris, L. E., Heinemann, M., Lapeña Jr, J. F., … Habibzadeh, F. (2023). Recomendaciones de WAME sobre chatbots e inteligencia artificial generativa en relación con las publicaciones académicas. *Colombia Médica*, *54*(3). https://doi.org/10.25100/cm.v54i3.5868