

PRINCIPIOS ÉTICOS EN EL PROCESO DE INVESTIGACIÓN CON EL USO DE INTELIGENCIA ARTIFICIAL

YAMIRENA CHÁVEZ IBARRA¹

yamich37@yahoo.es

ORCID: 0009-0004-8101-7863

Institución Educativa de San Miguel
Municipio de Hatillo de Loba

Recibido: 12/01/2026

Revisado: 14/02/2026

Aprobado: 18/03/2026

RESUMEN

El avance tecnológico y la adopción creciente de herramientas de inteligencia artificial (IA) transformaron los procesos de producción, análisis y divulgación del conocimiento, pero también generaron riesgos relacionados con la integridad científica, la autoría, los sesgos y la protección de datos. El objetivo del artículo fue analizar los principios éticos que deben orientar al investigador en el uso de la IA y de los Modelos de Lenguaje de Gran Escala (LLM) durante el proceso investigativo. Se realizó una revisión documental sistematizada de 19 fuentes académicas, institucionales y normativas, publicadas principalmente entre 2022 y 2026, localizadas en Scopus, Web of Science, PubMed, Google Scholar, Dialnet, Redalyc y SciELO. Se incluyeron documentos pertinentes, revisados por pares o emitidos por organismos reconocidos, y se excluyeron materiales divulgativos sin respaldo científico. La información se organizó mediante una matriz de análisis documental y se examinó de forma temática y comparativa. Los resultados permitieron identificar como principios centrales la transparencia, la responsabilidad, la autonomía, la equidad y justicia, la privacidad, la protección de datos y el cumplimiento normativo. Asimismo, se estableció que la supervisión humana, la trazabilidad y la declaración del uso de IA constituyen condiciones indispensables para preservar la validez y la reproducibilidad de los hallazgos. Se concluyó que la IA representa un recurso de apoyo, mientras que la autoría, la verificación y la responsabilidad ética permanecen en el investigador.

Palabras clave: ética en investigación, inteligencia artificial, transparencia académica, responsabilidad científica, modelos de lenguaje de gran escala.

¹ Ingeniera de Sistemas, Politécnico Grancolombiano, Bogotá-Colombia; McS en Informática Educativa, Universidad Rafael Belloso Chazín, Maracaibo-Venezuela

ETHICAL PRINCIPLES IN THE RESEARCH PROCESS USING ARTIFICIAL INTELLIGENCE

ABSTRACT

Technological advances and the growing adoption of artificial intelligence (AI) tools transformed the production, analysis, and dissemination of knowledge, while also creating risks related to scientific integrity, authorship, bias, and data protection. The objective of this article was to analyze the ethical principles that should guide researchers in the use of AI and Large Language Models (LLMs) throughout the research process. A systematized documentary review of 19 academic, institutional, and regulatory sources, mainly published between 2022 and 2026, was conducted. The sources were retrieved from Scopus, Web of Science, PubMed, Google Scholar, Dialnet, Redalyc, and SciELO. Relevant peer-reviewed documents and materials issued by recognized organizations were included, whereas non-scholarly publications without scientific support were excluded. The information was organized in a documentary analysis matrix and examined through thematic and comparative analysis. The findings identified transparency, responsibility, autonomy, fairness and justice, privacy, data protection, and regulatory compliance as the central principles. Human oversight, traceability, and disclosure of AI use were also found to be essential conditions for preserving the validity and reproducibility of research findings. It was concluded that AI is a supporting resource, while authorship, verification, and ethical responsibility remain with the researcher.

Keywords: research ethics, artificial intelligence, academic transparency, scientific responsibility, large language models.

INTRODUCCIÓN

La incorporación de la inteligencia artificial (IA) y de los Modelos de Lenguaje de Gran Escala (LLM) transformó de manera significativa las actividades de búsqueda, análisis, escritura y divulgación científica. Watkins (2024) y Hosseini, Resnik y Holmes (2023), reconocen que estas herramientas pueden mejorar la eficiencia del trabajo académico, pero advierten que su uso sin criterios explícitos puede comprometer la transparencia, la autoría y la integridad del manuscrito. De forma complementaria, Guaña-Moya y Chipuxi-Fajardo (2023) señalan que la automatización también plantea riesgos de sesgo, inequidad y exposición de información personal.

Aunque la literatura reciente ha formulado orientaciones generales, persiste un problema no resuelto: los investigadores no siempre disponen de criterios operativos uniformes para decidir qué tareas pueden delegarse, cómo verificar las salidas, cuándo declarar el uso de IA y qué medidas deben adoptarse frente a datos sensibles. Esta brecha puede favorecer referencias inexistentes, resultados no reproducibles, decisiones poco explicables y prácticas que debilitan la confianza pública en la ciencia. En consecuencia, la pregunta que orientó el artículo fue: ¿qué principios éticos deben guiar al investigador en el uso de la IA durante las distintas etapas del proceso investigativo?

El estudio del problema resultó relevante porque la responsabilidad científica no puede transferirse a una herramienta que carece de juicio moral y capacidad para responder por sus resultados. Paguay et al. (2024) destacan la necesidad de autocontrol y compromiso con las buenas prácticas, mientras que la UNESCO (2022) sitúa la dignidad humana, la equidad, la privacidad y la supervisión humana como fundamentos de una gobernanza responsable; analizar estos aportes permitió integrar perspectivas académicas, editoriales y normativas en un marco aplicable a la práctica investigativa.

El objetivo del artículo fue analizar los principios éticos que deben orientar el uso de la inteligencia artificial y de los LLM en la investigación científica, con énfasis en la transparencia, la responsabilidad, la autonomía, la equidad y justicia, la privacidad, la protección de datos y el cumplimiento normativo. Para alcanzar este objetivo, primero se describe el método de revisión y los criterios de selección; posteriormente se examinan los aportes de la literatura, se relacionan los principios con los riesgos presentes en cada etapa del proceso y se propone un protocolo operativo; finalmente, se presentan las conclusiones y las implicaciones institucionales.

METODOLOGÍA

El artículo se desarrolló mediante una revisión documental sistematizada con enfoque analítico-crítico, orientada a identificar, seleccionar, comparar e interpretar

fuentes académicas, institucionales y normativas relacionadas con el uso ético de la inteligencia artificial y de los Modelos de Lenguaje de Gran Escala (LLM) en la investigación. El corpus final estuvo conformado por 19 documentos, organizados en una matriz que permitió examinar el tipo de fuente, el ámbito de aplicación, el riesgo ético, el principio propuesto y la recomendación derivada para la práctica científica.

La búsqueda comprendió publicaciones difundidas principalmente entre 2022 y 2026, aunque se admitieron documentos anteriores cuando poseían carácter legal, fundacional o canónico, como la Ley Estatutaria 1581 de 2012 sobre protección de datos personales. Las consultas se realizaron en Scopus, Web of Science (WoS), PubMed, Google Scholar, Dialnet, Redalyc y SciELO; además, se revisaron documentos de la UNESCO, el Parlamento Europeo, el National Institute of Standards and Technology (NIST), el Committee on Publication Ethics (COPE), el International Committee of Medical Journal Editors (ICMJE) y la World Association of Medical Editors (WAME).

La selección se efectuó mediante criterios de inclusión y exclusión aplicados de manera conjunta. Se incluyeron artículos científicos, revisiones, ensayos académicos y documentos institucionales publicados principalmente entre 2022 y 2026 que abordaran de forma explícita la transparencia, la responsabilidad, la autonomía, la equidad y justicia, la privacidad, la protección de datos o el cumplimiento normativo; asimismo, se exigió que estuvieran indexados en bases de datos reconocidas o avalados por

organismos internacionales y que pudieran consultarse en español o inglés mediante acceso abierto o repositorios institucionales. Se excluyeron las publicaciones anteriores a 2022 que no tuvieran carácter legal, fundacional o canónico, los textos de divulgación general sin respaldo científico, los documentos ajenos a los contextos investigativo, educativo o científico y las publicaciones de acceso restringido que no podían recuperarse mediante repositorios institucionales.

La recuperación de la información se desarrolló en tres fases. Primero se realizó una búsqueda exploratoria mediante combinaciones de términos relacionados con ética, inteligencia artificial, LLM, integridad científica, privacidad y autoría; después se revisaron títulos, resúmenes y datos bibliográficos de acuerdo con los criterios definidos; por último, se efectuó la lectura completa y la depuración del corpus hasta consolidar las 19 fuentes. Los documentos primarios correspondieron a artículos científicos revisados por pares, las fuentes secundarias reunieron revisiones y ensayos de reflexión, y los materiales institucionales y normativos comprendieron recomendaciones, marcos regulatorios y políticas emitidas por organismos nacionales e internacionales.

La calidad, la fiabilidad y la validez de los documentos se evaluaron de forma integrada mediante el rigor metodológico, la claridad del argumento, la trazabilidad de las fuentes, la actualidad, la revisión por pares y el reconocimiento de la revista o institución responsable. Los documentos normativos también se examinaron según su vigencia y pertinencia para el tratamiento de datos, la autoría y la gestión del riesgo.

Este análisis crítico permitió valorar la consistencia de los hallazgos y de las propuestas éticas, así como reconocer convergencias, divergencias y vacíos en la literatura.

La información recuperada se organizó en una matriz de análisis documental que incluyó autor, año, tipo de fuente, ámbito de aplicación, riesgo ético identificado, principio propuesto y recomendación para la práctica investigativa. Los contenidos se codificaron de forma temática, se agruparon en categorías convergentes y se compararon de manera constante para relacionar los aportes académicos con las directrices editoriales y los marcos regulatorios aplicables. Debido al carácter cualitativo y documental del estudio, no se realizó metaanálisis; la validez interpretativa se fortaleció mediante la triangulación de artículos revisados por pares, orientaciones institucionales y documentos normativos.

DESARROLLO Y DISCUSIÓN

El desarrollo y la discusión se organizan en cuatro momentos articulados. El primero contrasta los aportes académicos, institucionales y normativos que fundamentan los principios éticos; el segundo examina la aplicación de la transparencia, la responsabilidad, la autonomía, la equidad y justicia, la privacidad, la protección de datos y el cumplimiento normativo; el tercero relaciona estos principios con los riesgos presentes en las distintas etapas del proceso investigativo; y el cuarto presenta un protocolo operativo junto con sus implicaciones institucionales y su aplicación en el

contexto colombiano. Esta secuencia permite avanzar desde la comprensión conceptual hacia criterios concretos para orientar el uso responsable de la inteligencia artificial en la investigación.

Watkins (2024), en el artículo “Guidance for researchers and peer-reviewers on the ethical use of Large Language Models (LLMs) in scientific research workflows”, sostiene que el investigador debe documentar el papel de los LLM en cada etapa del flujo científico. Su planteamiento ubica la transparencia como condición para reconstruir las decisiones, evaluar la contribución humana y detectar errores antes de que se incorporen a los resultados.

Hosseini, Resnik y Holmes (2023), en “The ethics of disclosing the use of artificial intelligence tools in writing scholarly manuscripts”, analizan las políticas editoriales sobre IA y argumentan que la apertura, la honestidad y la declaración proporcional del uso protegen la integridad científica. Los autores advierten que una prohibición absoluta puede fomentar usos ocultos, por lo que resultan preferibles reglas verificables y responsabilidades claramente asignadas.

Martínez González (2023), en “Uso responsable de la inteligencia artificial en estudiantes universitarios: una mirada rectoria”, explica que el uso orientado únicamente a ahorrar tiempo puede disminuir el análisis crítico y la creatividad. Este aporte evidencia que la competencia técnica debe acompañarse de formación ética y

de una supervisión que mantenga el equilibrio entre la inteligencia humana y la asistencia automatizada.

Guaña-Moya y Chipuxi-Fajardo (2023), en “Impacto de la inteligencia artificial en la ética y la privacidad de los datos”, identifican la transparencia, la responsabilidad, la equidad, el respeto de los derechos fundamentales y la privacidad como ejes de una implementación responsable. En la misma dirección, Paguay et al. (2024) destacan que el autocontrol, las buenas prácticas y la responsabilidad personal son necesarios para evitar que la IA sustituya el juicio crítico del investigador.

Aparicio-Gómez (2024), en “Consideraciones éticas para el uso académico de sistemas de inteligencia artificial”, plantea que la autonomía intelectual, la igualdad de acceso y la explicabilidad de los procedimientos son indispensables para preservar la calidad del conocimiento. Esta perspectiva vincula la ética con la capacidad del investigador para revisar, rechazar o corregir las respuestas automatizadas.

En el ámbito institucional, la UNESCO (2022) fundamenta la gobernanza de la IA en la dignidad, los derechos humanos, la equidad y la supervisión humana. El Parlamento Europeo (2024) adopta un enfoque regulatorio basado en niveles de riesgo, mientras que Breceda Pérez (2025) propone lineamientos aplicables a instituciones latinoamericanas. La Tabla 1 sintetiza las fuentes que sustentan cada principio y establece la relación con los apartados desarrollados a continuación.

Tabla 1. *Resumen de revisión bibliográfica por principio ético*

PRINCIPIO	REFERENCIAS CLAVE
Transparencia	Watkins (2024); Hosseini et al. (2023); UNESCO (2022)
Responsabilidad	Resnik y Hosseini (2024); Martínez González (2023); NIST (2024)
Autonomía	Suárez-Muñoz y Tejeda (2024); Aparicio-Gómez (2024); UNESCO (2022)
Equidad y justicia	UNESCO (2022); Aparicio-Gómez (2024); NIST (2024)
Privacidad	Guaña-Moya y Chipuxi-Fajardo (2023); UNESCO (2022); Parlamento Europeo (2024)
Protección de datos	Congreso de Colombia (2012); Parlamento Europeo (2024); NIST (2024)
Cumplimiento normativo	Parlamento Europeo (2024); Breceda Pérez (2025); Melgarejo Riveros y Rodas (2024)

Nota. Elaboración propia a partir de las fuentes incluidas en la revisión documental.

PRINCIPIOS ÉTICOS EN EL PROCESO DE INVESTIGACIÓN CON EL USO DE INTELIGENCIA ARTIFICIAL (IA)

TRANSPARENCIA EN EL PROCESO INVESTIGATIVO

La transparencia exige explicar qué herramienta se utilizó, con qué finalidad, en qué etapa y bajo qué controles. Watkins (2024) propone documentar la intervención de los LLM en el flujo de trabajo; Hosseini et al. (2023) relacionan la declaración con la honestidad y la correcta atribución; y la UNESCO (2022) exige que los sistemas sean comprensibles y sometidos a supervisión humana. En consecuencia, toda salida automatizada debe verificarse antes de incorporarse al análisis, y el investigador debe conservar evidencia suficiente para reconstruir las decisiones adoptadas.

RESPONSABILIDAD

La responsabilidad implica que la autoría y las consecuencias del estudio permanecen en las personas que lo suscriben. Resnik y Hosseini (2024) advierten que los sistemas de IA no pueden responder por errores, conflictos de interés o daños derivados de sus salidas. Por ello, el investigador debe reconocer las limitaciones del modelo, evaluar los sesgos de los datos y establecer controles proporcionales al riesgo de la tarea.

Martínez González (2023) complementa esta perspectiva al señalar que la formación del usuario es decisiva para evitar una dependencia acrítica de la herramienta. El NIST (2024) agrega que la gestión del riesgo debe contemplar confabulaciones, sesgos, privacidad y dependencia excesiva. Estos aportes permiten interpretar la responsabilidad como una práctica continua de verificación, documentación y corrección, y no como una declaración posterior al uso de la IA.

TOMA DE DECISIONES

La toma de decisiones no puede basarse exclusivamente en resultados automatizados. Suárez-Muñoz y Tejeda (2024) sostienen que la IA no posee autonomía comparable con la humana; por tanto, la selección de fuentes, la interpretación de datos y la aceptación de una conclusión deben permanecer bajo control del investigador. Esta obligación se intensifica cuando las decisiones pueden afectar a participantes, comunidades o instituciones.

Antes de adoptar una salida, el investigador debe contrastarla con fuentes primarias, revisar supuestos y estimar las consecuencias de un posible error. El criterio humano no se limita a aprobar el resultado final: también define la pregunta, selecciona el método, identifica los límites y justifica por qué una herramienta resulta pertinente para la tarea.

AUTONOMÍA

La autonomía se entiende como la conservación del criterio propio frente a las recomendaciones de un sistema automatizado. Suárez-Muñoz y Tejeda (2024) rechazan la idea de una autonomía moral de la IA; Aparicio-Gómez (2024) destaca la necesidad de proteger el pensamiento crítico; y la UNESCO (2022) establece que la supervisión humana debe permitir intervenir, corregir o detener el proceso. En consecuencia, delegar una tarea no equivale a delegar la decisión ni la responsabilidad intelectual.

EQUIDAD Y JUSTICIA

La equidad y la justicia exigen que la IA no reproduzca desigualdades presentes en los datos ni restrinja el acceso al conocimiento. Aparicio-Gómez (2024) relaciona la igualdad de acceso con la autonomía académica; la UNESCO (2022) demanda prevenir la discriminación; y el NIST (2024) recomienda evaluar sesgos y efectos diferenciados. Para aplicar este principio, el investigador debe examinar la

representatividad de los datos, comparar resultados entre grupos y declarar las limitaciones que puedan afectar la generalización de los hallazgos.

PRIVACIDAD Y PROTECCIÓN DE DATOS

La privacidad y la protección de datos son esenciales cuando la IA procesa entrevistas, registros escolares, información laboral, imágenes o datos de salud. Guaña-Moya y Chipuxi-Fajardo (2023) advierten que la recopilación automatizada puede exponer información sensible; la UNESCO (2022) exige respetar la dignidad y la finalidad del tratamiento; y el Parlamento Europeo (2024) adopta obligaciones diferenciadas según el nivel de riesgo.

En Colombia, la Ley Estatutaria 1581 de 2012 constituye el marco transversal para la protección de datos personales y, en este primer ámbito, orienta el manejo general de la información durante todo el ciclo investigativo mediante los principios de legalidad, finalidad, veracidad, acceso restringido, seguridad y confidencialidad. Su aplicación exige reducir la exposición de datos identificables, conocer las condiciones de almacenamiento, reutilización y eliminación de las plataformas empleadas, e incorporar medidas como la seudonimización, la minimización y el control de accesos desde la planificación del estudio. El NIST (2024) complementa este enfoque al recomendar la documentación de los riesgos de privacidad y la revisión continua de los controles, de modo que la protección no se limite a retirar nombres, sino que considere la posibilidad de reidentificación y el ciclo de vida completo de los datos.

CUMPLIMIENTO NORMATIVO

El cumplimiento normativo exige articular las obligaciones legales con las reglas editoriales e institucionales. El Parlamento Europeo (2024) establece un marco basado en riesgos; Breceda Pérez (2025) propone lineamientos para regular la IA generativa en instituciones académicas; y Melgarejo Riveros y Rodas (2024) destacan la necesidad de una regulación jurídica que acompañe la transformación de la investigación científica. En Colombia, estas orientaciones deben integrarse con la Ley 1581 de 2012 y con las políticas de cada institución.

La convergencia de los principios analizados demuestra que la eficiencia tecnológica no es un criterio suficiente para justificar el uso de IA. Una herramienta puede reducir tiempos y ampliar la capacidad de procesamiento, pero su intervención solo resulta legítima cuando mantiene la trazabilidad, protege a las personas y permite verificar las decisiones. La calidad metodológica depende, por tanto, de la relación entre competencia técnica y juicio ético.

El uso negligente de la IA puede afectar la dignidad, la intimidad, la reputación del investigador y la confianza en la ciencia. Estos efectos se agravan cuando se aceptan resultados sin verificación, se incorporan referencias inexistentes o se procesan datos sensibles en servicios no autorizados. Por ello, la supervisión debe ser permanente y proporcional a la influencia de la herramienta en los resultados. En síntesis, la inteligencia artificial debe concebirse como un recurso de apoyo y no como

sustituto de la capacidad analítica. El investigador conserva la obligación de formular el problema, seleccionar las fuentes, interpretar los datos, reconocer incertidumbres y responder por las conclusiones. Esta postura sustenta los apartados siguientes sobre integridad, trazabilidad, consentimiento, supervisión humana y riesgos en las etapas del proceso.

INTEGRIDAD CIENTÍFICA, AUTORÍA Y PROPIEDAD INTELECTUAL

La integridad científica exige que la utilización de IA no altere la atribución real de las ideas, los datos y las decisiones metodológicas. Un sistema generativo puede proponer estructuras, reformular párrafos o producir fragmentos de código, pero no puede asumir responsabilidad por la originalidad, la exactitud ni las consecuencias de un manuscrito. COPE (2023) establece que las herramientas de IA no cumplen los criterios de autoría, porque no pueden responder por el trabajo, declarar conflictos de interés ni suscribir acuerdos editoriales. Por tanto, incluir una aplicación como autora constituye una atribución incorrecta y desplaza indebidamente la responsabilidad humana.

La transparencia sobre el uso de IA debe ser proporcional a su incidencia en el estudio. Cuando la herramienta se limita a corregir ortografía, su declaración puede ser breve; si interviene en la síntesis de literatura, la generación de código, la traducción, la clasificación de datos o la elaboración de figuras, resulta necesario indicar la herramienta, la versión disponible, la finalidad de uso y los mecanismos de verificación

aplicados. Hosseini et al. (2023) sostienen que la divulgación permite valorar con mayor precisión la contribución humana y evita que la asistencia automatizada se confunda con producción intelectual autónoma.

También debe prevenirse el plagio involuntario y la vulneración de derechos de autor. Los LLMs (Modelos de Lenguaje de gran escala), generan respuestas a partir de regularidades estadísticas y pueden reproducir expresiones cercanas a materiales existentes sin identificar su procedencia. En consecuencia, el investigador debe verificar similitudes, localizar las fuentes originales y citar únicamente documentos consultados de manera directa. Las referencias propuestas por la IA no deben incorporarse sin comprobar autoría, título, revista, volumen, páginas y enlace permanente. La responsabilidad por una cita inexistente o por una paráfrasis inadecuada recae exclusivamente en quien firma el trabajo.

VERACIDAD, TRAZABILIDAD Y REPRODUCIBILIDAD

La veracidad constituye una obligación central porque los sistemas generativos pueden producir afirmaciones plausibles, aunque incorrectas, incompletas o descontextualizadas. Este fenómeno resulta especialmente riesgoso cuando el investigador utiliza respuestas automatizadas como si fueran evidencia. La IA no reemplaza la consulta de artículos, bases de datos, protocolos ni documentos legales. Cada dato relevante debe contrastarse con una fuente primaria o institucional y toda inferencia debe distinguirse de los hechos documentados. El ICMJE (2026) advierte

que los autores deben asegurar la ausencia de plagio y verificar la atribución de los materiales generados o modificados mediante IA.

La trazabilidad implica conservar un registro suficiente de las interacciones que influyeron en el resultado. Este registro puede incluir fecha, nombre y versión de la herramienta, instrucciones empleadas, archivos suministrados, respuestas relevantes, correcciones humanas y decisión final. No se trata de publicar indiscriminadamente todas las conversaciones, sino de mantener evidencia que permita explicar cómo se obtuvo un resultado y qué controles se aplicaron. Watkins (2024) propone documentar el papel de los LLM a lo largo del flujo de trabajo científico, pues la trazabilidad facilita la auditoría, la revisión por pares y la identificación de errores.

La reproducibilidad exige además separar los procedimientos deterministas de las respuestas que pueden variar entre ejecuciones. Cuando la IA se utiliza para programar, depurar código o apoyar un análisis, el investigador debe conservar el código finalmente ejecutado, las versiones de las bibliotecas, los parámetros y los datos de entrada. El marco AI RMF del NIST (2023) promueve una gestión continua del riesgo basada en gobernar, mapear, medir y gestionar. Trasladado a la investigación, este enfoque demanda anticipar fallos, evaluar su impacto, implementar controles y revisar periódicamente si la herramienta sigue siendo adecuada para la finalidad declarada.

CONSENTIMIENTO INFORMADO Y TRATAMIENTO DE DATOS SENSIBLES

El consentimiento informado adquiere una función específica cuando el uso de inteligencia artificial modifica la forma en que serán tratados los datos obtenidos durante la investigación. Procesar entrevistas, historias clínicas, imágenes, registros escolares o información laboral mediante plataformas externas puede implicar que la información sea gestionada por terceros y bajo condiciones diferentes de las previstas inicialmente. Por esta razón, antes de incorporar una herramienta de IA, el investigador debe establecer si el tratamiento previsto corresponde con la finalidad informada a los participantes y si las condiciones de uso de la plataforma son compatibles con las garantías ofrecidas durante la recolección de la información.

La Ley Estatutaria 1581 de 2012 también resulta aplicable en este segundo ámbito, aunque su función se diferencia de la protección general de datos explicada anteriormente. Mientras el apartado de privacidad se concentra en las medidas de seguridad, confidencialidad y reducción de la exposición de la información, en el consentimiento informado la atención se dirige a que la persona conozca la finalidad del tratamiento y las condiciones bajo las cuales sus datos podrán ser procesados. Cuando el consentimiento original no contempla el uso de servicios de inteligencia artificial o este introduce condiciones distintas para el manejo de la información, el investigador debe evaluar la necesidad de una autorización adicional o recurrir a un entorno institucional controlado que mantenga las condiciones previamente informadas.

La aplicación de estas garantías requiere mayor cuidado cuando se trabaja con información sensible o con poblaciones que demandan protección reforzada. El Congreso de Colombia (2012) establece consideraciones especiales para los datos de niños, niñas y adolescentes, mientras que la UNESCO (2022) recomienda salvaguardas proporcionales cuando la inteligencia artificial puede afectar a personas o grupos vulnerables. En estos casos, la conveniencia tecnológica no justifica ampliar el uso de la información más allá de lo autorizado, por lo que deben mantenerse medidas de minimización, seudonimización, control de accesos y eliminación segura, además de consultar al comité de ética cuando la incorporación de IA modifique los riesgos inicialmente aprobados. Así, privacidad y consentimiento se articulan como obligaciones complementarias, pues la primera protege el dato durante su ciclo de vida y el segundo delimita las condiciones bajo las cuales su tratamiento resulta aceptable dentro del estudio.

SUPERVISIÓN HUMANA Y COMPETENCIA ÉTICO-DIGITAL

La supervisión humana no consiste en aceptar una respuesta automática y revisarla de manera superficial; implica conservar autoridad real para detener el proceso, modificar el procedimiento y rechazar resultados que no sean explicables o verificables. La UNESCO (2023) plantea una visión centrada en el ser humano para la IA generativa en educación e investigación; desde esta perspectiva, la herramienta debe ampliar capacidades sin sustituir la deliberación, la creatividad y la

responsabilidad moral. El investigador debe comprender qué tareas delega, por qué las delega y cuáles son los límites de la salida obtenida.

Esta exigencia supone desarrollar competencia ético-digital. No basta con formular instrucciones eficientes; se requiere reconocer sesgos, evaluar la calidad de las fuentes, identificar riesgos de privacidad y comprender las restricciones de propiedad intelectual. La formación debe integrar elementos metodológicos, jurídicos y disciplinares, porque una respuesta técnicamente correcta puede ser epistémicamente débil o éticamente inaceptable. La revisión por un experto del área sigue siendo indispensable cuando la IA interviene en contenidos especializados o en decisiones con efectos sobre personas.

La supervisión también debe ser colectiva. Las instituciones han de definir políticas claras, canales de consulta, mecanismos de reporte y criterios diferenciados según el nivel de riesgo. Un uso permitido para lluvia de ideas puede ser inaceptable para evaluar participantes o revisar un manuscrito confidencial. Por ello, la alfabetización en IA debe acompañarse de una cultura de integridad en la que declarar el uso de herramientas no sea percibido como una falta, sino como una práctica de transparencia y rendición de cuentas.

RIESGOS ÉTICOS EN LAS ETAPAS DEL PROCESO INVESTIGATIVO

Los riesgos no se concentran únicamente en la redacción final. La IA puede influir desde la formulación del problema hasta la publicación, y cada etapa exige controles específicos. Analizar el ciclo completo permite evitar que una decisión temprana, aparentemente menor, se convierta en una fuente de sesgo o invalidez difícil de detectar al final del estudio.

FORMULACIÓN DEL PROBLEMA Y REVISIÓN DE LITERATURA

Durante la delimitación del problema, la IA puede ayudar a explorar conceptos y generar términos de búsqueda, pero también puede orientar la investigación hacia temas más representados en sus datos de entrenamiento. Para reducir este sesgo, las preguntas generadas deben contrastarse con necesidades reales, antecedentes verificables y conocimiento del contexto. En la revisión de literatura, los resúmenes automáticos deben utilizarse como apoyo preliminar y no como sustituto de la lectura crítica. La selección de estudios debe obedecer a criterios explícitos, y las citas deben provenir de documentos localizados y leídos por el investigador.

RECOLECCIÓN, PROCESAMIENTO Y ANÁLISIS DE DATOS

En la recolección y el análisis, los principales riesgos son la exposición de información, la alteración silenciosa de los datos y la aceptación de patrones espurios. El investigador debe trabajar con copias controladas, documentar las transformaciones

y validar los resultados mediante procedimientos estadísticos o cualitativos reconocidos. Si una aplicación clasifica respuestas, codifica entrevistas o detecta categorías, una muestra debe revisarse manualmente y las discrepancias deben analizarse. La eficiencia no puede obtenerse a costa de perder el vínculo entre los datos originales y las conclusiones.

REDACCIÓN, REVISIÓN POR PARES Y PUBLICACIÓN

En la redacción, la IA puede mejorar claridad y coherencia, pero existe riesgo de homogeneizar la voz académica, introducir afirmaciones no sustentadas u ocultar vacíos argumentativos. Antes de enviar un manuscrito, cada autor debe revisar la totalidad del texto y poder defenderlo. En la revisión por pares, la confidencialidad adquiere especial importancia: el ICMJE (2026) señala que editores y revisores no deben cargar manuscritos en sistemas donde no pueda garantizarse la confidencialidad sin autorización. WAME (2023) coincide en que el uso de chatbots debe declararse y someterse a políticas editoriales explícitas.

PROTOCOLO PARA EL USO ÉTICO DE IA EN INVESTIGACIÓN

A partir de los principios analizados se propone un protocolo operativo adaptable a proyectos documentales, educativos, sociales, tecnológicos o de salud. La Tabla 2 presenta el protocolo en seis fases articuladas —planificación, selección, interacción, verificación, autoría y conservación— y relaciona cada fase con la acción esperada del

investigador y la evidencia mínima que debe conservarse. Esta organización convierte los principios generales en decisiones observables y verificables, cuyo nivel de documentación, validación y supervisión debe aumentar cuando los datos sean más sensibles o la IA tenga mayor influencia en los resultados.

Tabla 2. *Protocolo operativo para el uso ético de inteligencia artificial en investigación*

FASE	ACCIÓN DEL INVESTIGADOR	EVIDENCIA MÍNIMA
1. Planificación	Definir finalidad, nivel de riesgo, datos que podrán utilizarse y tareas que no se delegarán.	Registro de riesgos y decisión de uso.
2. Selección	Revisar privacidad, condiciones del servicio, versión, capacidades y limitaciones de la herramienta.	Ficha técnica o nota metodológica.
3. Interacción	Evitar datos identificables, formular instrucciones delimitadas y conservar las interacciones relevantes.	Prompts, fecha, versión y archivos autorizados.
4. Verificación	Contrastar hechos y referencias, revisar sesgos, reproducir análisis y validar con criterios disciplinares.	Lista de comprobación y fuentes verificadas.
5. Autoría	Reescribir críticamente, asumir responsabilidad integral y declarar la contribución de la IA.	Declaración de uso en el manuscrito.
6. Conservación	Archivar versiones, código, decisiones y resultados de control conforme a la política institucional.	Repositorio o expediente de trazabilidad.

Nota. Elaboración propia a partir de los principios éticos y las fuentes examinadas en la revisión documental.

El protocolo busca asegurar que la declaración de uso no sea un trámite aislado. La transparencia debe estar respaldada por evidencias que permitan reconstruir el proceso y evaluar si la herramienta fue pertinente. El NIST (2024) destaca riesgos específicos de la IA generativa, entre ellos la confabulación, la privacidad, el sesgo y la dependencia excesiva de las salidas. Por ello, la verificación debe combinar controles automáticos con revisión humana experta y no limitarse a comprobar la fluidez del texto.

Una declaración adecuada debe identificar las etapas en las que se utilizó IA, la finalidad y el tipo de intervención. No es necesario atribuir a la herramienta decisiones

que pertenecen al investigador; por el contrario, debe quedar claro que la selección de fuentes, el diseño, el análisis y las conclusiones fueron evaluados y aprobados por personas responsables. La documentación interna puede ser más detallada que la declaración publicada, especialmente cuando contiene instrucciones, código o información protegida.

IMPLICACIONES INSTITUCIONALES Y CONTEXTO COLOMBIANO

La ética individual requiere apoyo institucional. Las universidades, centros de investigación y revistas deben establecer políticas comprensibles que diferencien usos permitidos, condicionados y prohibidos. Estas políticas deben armonizar integridad académica, protección de datos, seguridad de la información y derechos de autor. También deben prever procedimientos para actualizar lineamientos, dado que las capacidades y condiciones de las plataformas cambian con rapidez. Una norma demasiado general deja al investigador sin criterios, mientras que una prohibición absoluta puede fomentar usos ocultos y dificultar la formación.

En el contexto colombiano, las obligaciones derivadas de la Ley 1581 de 2012 deben traducirse institucionalmente en procedimientos que orienten al investigador antes de utilizar servicios externos con información personal. Los comités de ética y las oficinas responsables de la protección de datos pueden apoyar la evaluación de los riesgos, revisar si el tratamiento previsto mantiene las condiciones informadas a los

participantes y definir alternativas seguras cuando se trabajen datos sensibles. Asimismo, las instituciones necesitan repositorios controlados, herramientas autorizadas y programas de capacitación que permitan aplicar de forma coherente las medidas de privacidad y los criterios de consentimiento desarrollados en los apartados anteriores, sin profundizar las brechas entre investigadores con distinto acceso tecnológico.

La gobernanza institucional debe combinar prevención y acompañamiento. Además de sancionar conductas fraudulentas, debe ofrecer asesoría para declarar el uso de IA, revisar contratos de tratamiento de datos y resolver dudas sobre autoría o confidencialidad. Este enfoque fortalece una cultura en la que la innovación es compatible con la responsabilidad científica y permite que la adopción tecnológica responda a los objetivos de la investigación, en lugar de imponerlos.

CONCLUSIONES

El objetivo del artículo se cumplió al identificar que el uso ético de la inteligencia artificial en la investigación debe regirse por la transparencia, la responsabilidad, la autonomía, la equidad y justicia, la privacidad, la protección de datos y el cumplimiento normativo. Estos principios no funcionan de manera aislada: en conjunto garantizan que la eficiencia tecnológica no desplace la validez metodológica, la dignidad de las personas ni la responsabilidad científica.

Los autores revisados aportaron dimensiones complementarias. Watkins (2024) y Hosseini et al. (2023) fundamentaron la transparencia y la declaración del uso; Resnik y Hosseini (2024) reafirmaron la responsabilidad humana; Suárez-Muñoz y Tejeda (2024) delimitaron la autonomía; Guaña-Moya y Chipuxi-Fajardo (2023) desarrollaron los riesgos de privacidad; y la UNESCO (2022), el NIST (2024) y el Parlamento Europeo (2024) integraron derechos humanos, gestión del riesgo y cumplimiento regulatorio. La convergencia de estos aportes permitió construir una visión analítica que supera la simple enumeración de normas.

El aporte central del artículo consistió en traducir los principios generales en un protocolo operativo de planificación, selección, interacción, verificación, autoría y conservación. Este protocolo facilita que el investigador documente la finalidad, el nivel de riesgo, las interacciones relevantes, las fuentes verificadas y la decisión final. De esta manera, la declaración del uso de IA deja de ser un requisito formal y se convierte en evidencia de trazabilidad y rendición de cuentas.

La responsabilidad ética debe evaluarse durante todo el ciclo investigativo. La formulación del problema, la revisión de literatura, la recolección y el análisis de datos, la redacción y la revisión por pares presentan riesgos diferentes y requieren controles específicos. La supervisión humana debe conservar la capacidad real de corregir, detener o rechazar una salida cuando no sea verificable, explicable o compatible con los derechos de los participantes.

El acompañamiento institucional constituye una condición decisiva para que los principios identificados se apliquen de manera estable en la práctica científica. Las universidades, los centros de investigación y las revistas deben mantener políticas actualizadas, ofrecer formación ético-digital, disponer de infraestructura segura y fortalecer la orientación de los comités de ética y de protección de datos, de modo que el investigador cuente con criterios claros para seleccionar herramientas, proteger la información, verificar resultados y declarar el uso de IA. Estas medidas también deben asignar responsables, establecer mecanismos de seguimiento y definir rutas de atención frente a errores, sesgos o incidentes de confidencialidad, además de procurar un acceso equitativo a recursos y capacitación. Bajo estas condiciones, la inteligencia artificial puede contribuir a una ciencia más eficiente y accesible sin desplazar el juicio crítico ni la responsabilidad humana, aunque la evolución acelerada de las tecnologías exige revisar periódicamente los lineamientos y ajustarlos a los nuevos riesgos y capacidades.

REFERENCIAS

- Aparicio-Gómez, O.-Y. (2024). Consideraciones éticas para el uso académico de sistemas de inteligencia artificial. *Revista Internacional de Filosofía Teórica y Práctica*, 4(1), 175–198. <https://doi.org/10.51660/RIFTP.V4I1.95>
- Breceda Pérez, J. A. (2025). Hacia la regulación ética del uso de inteligencia artificial en la investigación científica: lineamientos desde la Universidad Autónoma de Ciudad Juárez. *Revista de Bioética y Derecho*, (64), 81–97. <https://doi.org/10.1344/rbd2025.64.48719>
- Committee on Publication Ethics [COPE]. (2023). Authorship and AI tools. <https://publicationethics.org/guidance/cope-position/authorship-and-ai-tools>
- Congreso de Colombia. (2012, 17 de octubre). Ley Estatutaria 1581 de 2012, por la cual se dictan disposiciones generales para la protección de datos personales. Diario Oficial No. 48.587. <https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=49981>.
- Guaña-Moya, J., & Chipuxi-Fajardo, L. (2023). Impacto de la inteligencia artificial en la ética y la privacidad de los datos. *RECIAMUC*, 7(1), 923-930. [https://doi.org/10.26820/reciamuc/7.\(1\).enero.2023.923-930](https://doi.org/10.26820/reciamuc/7.(1).enero.2023.923-930)
- Hosseini, M., Resnik, D. B., & Holmes, K. (2023). The ethics of disclosing the use of artificial intelligence tools in writing scholarly manuscripts. *Research Ethics*, 19(4), 449–465. <https://doi.org/10.1177/17470161231180449>
- International Committee of Medical Journal Editors [ICMJE]. (2026). Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals. <https://www.icmje.org/recommendations/>
- Martínez González, M. A. (2023). Uso responsable de la inteligencia artificial en estudiantes universitarios: Una mirada rectoria. *Revista Boletín Redipe*, 12(9), 172–178. <https://doi.org/10.36260/rbr.v12i9.2008>
- Miao, F., & Holmes, W. (2023). Guidance for generative AI in education and research. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000386693>
- Melgarejo Riveros, V. M., & Rodas R., L. A. (2024). El papel de la Inteligencia artificial en los procesos de investigación científica y la necesidad de la regulación jurídica. *Arandu Poty*, 3(2), 18–23.

REVISIÓN BIBLIOGRÁFICA

- National Institute of Standards and Technology [NIST]. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0) (NIST AI 100-1). U.S. Department of Commerce. <https://doi.org/10.6028/NIST.AI.100-1>
- National Institute of Standards and Technology [NIST]. (2024). Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile (NIST AI 600-1). U.S. Department of Commerce. <https://doi.org/10.6028/NIST.AI.600-1>
- Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura [UNESCO]. (2022). Recomendación sobre la ética de la inteligencia artificial. UNESCO.
- Paguay-Simbaña, M. Y., Jimenez-Abad, D., Quiliguango-Lanchimba, V. F., Maynaguez-Canacuan, M. P., de los Ángeles Coello-García, C., & Coello-Ortiz, S. M. (2024). La ética en el uso de la inteligencia artificial en los procesos educativos. *Revista científica retos de la ciencia*, 1(4), 145-158.
- Parlamento Europeo. (2024). *Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de Inteligencia Artificial — AI Act)*. Diario Oficial de la Unión Europea. <https://artificialintelligenceact.eu/>
- Resnik, D. B., & Hosseini, M. (2024). The ethics of using artificial intelligence in scientific research: new guidance needed for a new tool. *AI and Ethics*, 5(2), 1499–1521. <https://doi.org/10.1007/s43681-024-00493-8>
- Suárez-Muñoz, F., & Tejeda, A. A. B. (2024). Autonomía de la inteligencia artificial. Una reflexión crítica desde la filosofía kantiana. *SUMMA*, 6(1), 1-11. <https://doi.org/10.47666/summa.6.1.9>
- Watkins, R. (2024). Guidance for researchers and peer-reviewers on the ethical use of large language models (LLMs) in scientific research workflows. *AI and Ethics*, 4, 969–974. <https://doi.org/10.1007/s43681-023-00294-5>
- World Association of Medical Editors [WAME]. (2023). Chatbots, generative AI, and scholarly manuscripts: WAME recommendations on chatbots and generative artificial intelligence in relation to scholarly publications. <https://wame.org/page3.php?id=106>