Masferrer Investiga
Revista Científica de la Universidad Salvadoreña Alberto Masferrer - USAM

Revista Masferrer Investiga, Vol. 15, N° 3, [pp. 41-57]
Septiembre, 2025
ISSN: 2079-9020 | e-ISSN: 2079-9039
DOI: https://doi.org/10.65880/ba5myc15

IA responsable en el ciclo de la investigación: revisión de alcance PRISMA-ScR (2024–2025)

Responsible AI in the research lifecycle: a PRISMA-ScR scoping review (2024–2025)

Información del articulo

Fecha de recepción: mayo de 2025

Fecha de aceptación: julio de 2025

Como citar este articulo (APA)

Oviedo-Argueta, A.; Vásquez, B. & Pérez de Martínez, J. (2025). IA responsable en el ciclo de la investigación: revisión de alcance PRISMA-ScR (2024–2025). *Masferrer Investiga, 15*(3), pp. 41-57. https://doi.org/10.65880/ba5myc15

Alberto Oviedo-Argueta
Universidad Salvadoreña Alberto Masferrer (El Salvador)
https://orcid.org/0009-0000-7474-2072
alberto.oviedo01@liveusam.edu.sv

Bruno Paúl Alexander Vásquez
Universidad Salvadoreña Alberto Masferrer (El Salvador)
https://orcid.org/0009-0006-5111-9230
bruno.vasquez01@liveusam.edu.sv

Johanna Beatriz Pérez de Martínez
Universidad Salvadoreña Alberto Masferrer (El Salvador)
https://orcid.org/0000-0002-4984-324X
johana.perez01@liveusam.edu.sv

Resumen

La irrupción de la IA generativa está reconfigurando el trabajo académico, equilibrando productividad e integridad. Esta revisión de alcance PRISMA-ScR, mapea la integración responsable de la IA a lo largo del ciclo de la investigación, búsqueda, escritura, revisión por pares, reporte y gobernanza editorial, y sintetiza prácticas aplicables en entornos universitarios y editoriales. La búsqueda en Scopus, limitada a inglés, acceso abierto y versión final entre 2024 y 2025, identificó 79 registros; tras el cribado temático se incluyeron 15 documentos que abarcan políticas y posicionamientos editoriales, estudios empíricos y metainvestigación, guías de reporte y propuestas metodológicas. El análisis mostró convergencia en principios clave: la IA no puede figurar como autora; la divulgación debe ser explícita e incluir herramienta, versión, propósito y, cuando corresponda, prompts; y los detectores presentan limitaciones y sesgos, por lo que no deben usarse como prueba única en decisiones editoriales. Las políticas evolucionan hacia esquemas escalonados con responsabilidad humana y resguardo de la confidencialidad, especialmente en revisión por pares, donde se desaconseja cargar manuscritos inéditos en servicios públicos y se acotan tareas asistibles por modelos bajo supervisión. Emergen listas de verificación y estándares de reporte que exigen identificar modelo y versión, semillas, prompts y criterios de evaluación, además de flujos con control de versiones para auditar contribuciones. Se documentan riesgos metodológicos, como la inestabilidad por repetición de prompts, lo que exige reproducibilidad y trazabilidad efectiva. Proponemos una arquitectura mínima de uso responsable que refuerza integridad, reproducibilidad y confianza en la comunicación científica.

Palabras clave: Inteligencia Artificial Generativa; Modelo De Lenguaje; Revisión Por Pares; Integridad Científica; Reproducibilidad.

Abstract

Generative artificial intelligence is reshaping scholarly work by expanding productivity while demanding rigorous safeguards for research integrity. This scoping review, aligned with PRISMA ScR, maps responsible integration of AI across the research

lifecycle, including search, writing, peer review, reporting, and editorial governance, and synthesizes practices for universities and journals. A structured Scopus search limited to English open access and final status during 2024 and 2025 identified seventy-nine records. After screening, fifteen documents were included, spanning editorial policies and statements, empirical and meta-research, reporting guidance, and methodological proposals. Analysis revealed convergence on core principles. Artificial intelligence cannot be credited as an author. Disclosure must be explicit and granular, specifying tool, model version, intended purpose, and prompts when relevant. Given known limitations and biases, AI detectors should not constitute sole evidence in editorial decisions. Policies are evolving toward tiered use with human accountability and confidentiality safeguards. In peer review, uploading unpublished manuscripts to public services is discouraged, and permitted tasks for language models are narrowly defined under supervision. Reporting checklists clearly require identification of model and version, seeds, prompts, and evaluation criteria. Version controlled workflows enable audit trails of AI and human contributions. Methodological risks, including instability from prompt repetition, underscore the need for reproducibility and robust traceability. We propose a minimal architecture for responsible adoption that implements disclosure matrices, human accountability, a limited role for detectors, supervised peer review protocols, reporting checklists, and transparent change logs. This architecture strengthens integrity, reproducibility, and trust across scholarly communication.

Keywords: Generative Artificial Intelligence; Large Language Model; Peer Review; Research Integrity; Reproducibility.

Introducción

La aceleración reciente en el desarrollo y adopción de sistemas de inteligencia artificial está transformando de forma transversal el ciclo de la investigación científica: desde la formulación de preguntas y la búsqueda de literatura, hasta la recolección/curación de datos, el análisis, la redacción y la evaluación por pares. Este despliegue simultáneo de modelos generativos, asistentes de revisión, motores de descubrimiento y herramientas de análisis plantea oportunidades inéditas de eficiencia y alcance, pero también riesgos concretos que comprometen la validez, la equidad y la reproducibilidad del conocimiento. En este contexto, la noción de IA responsable, que integra principios de transparencia, explicabilidad, privacidad, seguridad, justicia, rendición de cuentas e inclusión, ha pasado de ser un ideal normativo a una condición operativa para salvaguardar la integridad científica (Akinrinola et al., 2024; Carter et al., 2019; Istasy et al., 2021).

A pesar del auge de marcos y guías de "IA confiable", la evidencia empírica sobre cómo esos principios se aplican, o fallan, en las tareas cotidianas de investigación sigue fragmentada. Existen recomendaciones generales y casos de uso parciales, pero falta una síntesis que conecte los riesgos, salvaguardas y buenas prácticas etapa por etapa del ciclo de la investigación. El resultado es una brecha entre la formulación ética de alto nivel y las decisiones prácticas que toman investigadores, comités editoriales, revisores, bibliotecas/CRAI y agencias financiadoras al implementar o regular herramientas de IA (Li et a., 2023). Para atender esta brecha, realizamos una revisión de alcance (PRISMA-ScR) centrada en el período 2024–2025, con el objetivo de mapear sistemáticamente la literatura reciente sobre IA responsable a lo largo del ciclo de la investigación, identificar patrones, vacíos y tensiones, y proponer una taxonomía operativa que vincule tareas, riesgos y salvaguardas. Guiaron nuestro trabajo las siguientes preguntas:

- ¿Qué riesgos y fallos de la IA emergen de forma diferencial en cada etapa del ciclo (búsqueda, datos, análisis, escritura, evaluación, difusión y preservación)?
- ¿Qué salvaguardas técnicas, organizativas y procedimentales han mostrado mayor promesa para mitigar esos riesgos sin sacrificar productividad ni rigor?
- ¿Qué vacíos de evidencia persisten (métricas, auditorías, reportes de impacto, gobernanza), y dónde se concentran?
- ¿Qué recomendaciones accionables se pueden ofrecer a actores clave (investigadores, CRAI/bibliotecas, editores, financiadores, comités de ética) para operacionalizar la IA responsable con trazabilidad y auditoría?

Adoptamos el enfoque PRISMA-ScR por su idoneidad para campos emergentes y heterogéneos, priorizando estudios publicados entre 2024 y 2025 que describen aplicaciones, riesgos, marcos, auditorías o intervenciones sobre IA en procesos de investigación (Polanco & Mayorga, 2025). Tras la depuración, 15 artículos cumplieron criterios de elegibilidad y fueron analizados mediante extracción temática y síntesis narrativa. Esta selección cubre disciplinas y escenarios diversos, lo que permite observar regularidades y variaciones contextuales sin imponer una homogeneidad artificial (Teodoro et al., 2025).

Nuestro trabajo aporta cuatro contribuciones principales. Primero, proponemos una taxonomía etapa-tarea que alinea principios de IA responsable con riesgos específicos y controles verificables (por ejemplo, trazabilidad de prompts y datasets; pruebas de robustez y sesgo; políticas de autoría y divulgación de uso de IA; salvaguardas de privacidad; estándares de citación y verificación de fuentes). Segundo, sintetizamos prácticas prometedoras, técnicas y organizativas, que muestran factibilidad y valor añadido en entornos reales. Tercero, identificamos vacíos críticos: métricas comparables de impacto, evidencia sobre efectos en revisión por pares, costos de cumplimiento en grupos con menos recursos, e interoperabilidad de reportes entre instituciones (Echeverría et al., 2023; Hernández et al., 2023; Mayorga et al., 2023). Cuarto, derivamos un conjunto de recomendaciones operativas orientadas a decisiones, con niveles de madurez y responsables claros, para facilitar su adopción incremental en universidades y revistas (Lu et al., 2024).

El alcance de esta revisión es deliberadamente pragmático: no pretende dirimir debates filosóficos sobre la naturaleza de la inteligencia o la autonomía de los sistemas, sino traducir principios en controles implementables y en evidencias que permitan evaluar beneficios y riesgos con criterios compartidos. Consideramos que este enfoque es especialmente útil para actores que deben balancear innovación, cumplimiento y confianza pública, incluidos los CRAI que hoy articulan servicios, formación, curación de datos y soporte metodológico.

Materiales y método

El estudio adoptó una revisión de alcance para examinar de forma comprensiva la integración responsable de la IA generativa a lo largo del ciclo de la investigación. Esta elección metodológica responde al carácter emergente del campo y a la dispersión de la evidencia disponible, lo que hace pertinente un mapeo amplio antes que una síntesis evaluativa. La revisión se condujo conforme a la extensión PRISMA-ScR, que establece criterios mínimos y etapas secuenciadas para planificar, seleccionar, extraer y sintetizar de manera transparente.

Etapa 1. Planificación

Se delimitaron preguntas orientadas a: i) identificar riesgos y salvaguardas reportados por etapa del ciclo de la investigación; ii) caracterizar políticas, guías y estándares de uso y reporte de IA en publicación científica; y iii) localizar vacíos metodológicos y de gobernanza. Se diseñó una estrategia de búsqueda basada en términos controlados y sinónimos combinando "generative AI", "large language model", "LLM", "ChatGPT" con "research lifecycle", "scholarly publishing", "research integrity", "peer review", "authorship", "disclosure", "editorial policy", "guideline", "checklist", "PRISMA", "transparency". La búsqueda se ejecutó en Scopus por su cobertura multidisciplinaria y riqueza de metadatos. Para asegurar estabilidad editorial, se limitaron los resultados a documentos en inglés, acceso abierto, versión final, tipo artículo, publicados en 2024–2025. Se definieron exclusiones para trabajos técnico-clínicos sin vínculo con integridad o publicación, preprints, editoriales breves sin contenido normativo/metodológico, duplicados y documentos fuera de los filtros lingüísticos y de acceso (ver Tabla 1).

Tabla 1.

Criterios de inclusión y exclusión

Indicadores	Criterios de inclusión	Criterios de exclusión
Metodología / diseño	Artículos con enfoque cualitativo, cuantitativo o mixto, así como revisiones narrativas, de alcance o sistemáticas que analicen el uso responsable de IA en procesos de investigación, publicación o evaluación científica.	Estudios experimentales, preexperimentales o cuasiexperimentales enfocados en el desarrollo técnico de modelos de IA sin relación con la integridad, la ética o la publicación científica.
Fuente de publicación	Artículos publicados en revistas científicas revisadas por pares, capítulos de libro o actas de congreso con revisión editorial y relevancia académica.	Libros completos, informes institucionales, editoriales, comentarios de opinión, blogs, notas de prensa o cartas al editor sin revisión por pares.
Temática / sujeto	Artículos que aborden de forma explícita el uso de IA generativa o modelos de lenguaje (LLM) en el ciclo de la investigación, incluyendo escritura científica, revisión por pares, políticas editoriales, divulgación del uso de IA, reproducibilidad y estándares de reporte.	Trabajos sobre IA en contextos técnicos, clínicos o industriales sin relación con publicación científica, integridad académica o gobernanza editorial.
Periodo temporal considerado	Publicaciones comprendidas entre 2024 y 2025, priorizando artículos con versión final y acceso abierto publicados entre 2024 y 2025.	Publicaciones anteriores a 2024 o fuera del rango temporal establecido.
Idioma	Artículos publicados originalmente en inglés y disponibles en acceso abierto.	Artículos en otros idiomas o sin traducción disponible al inglés.

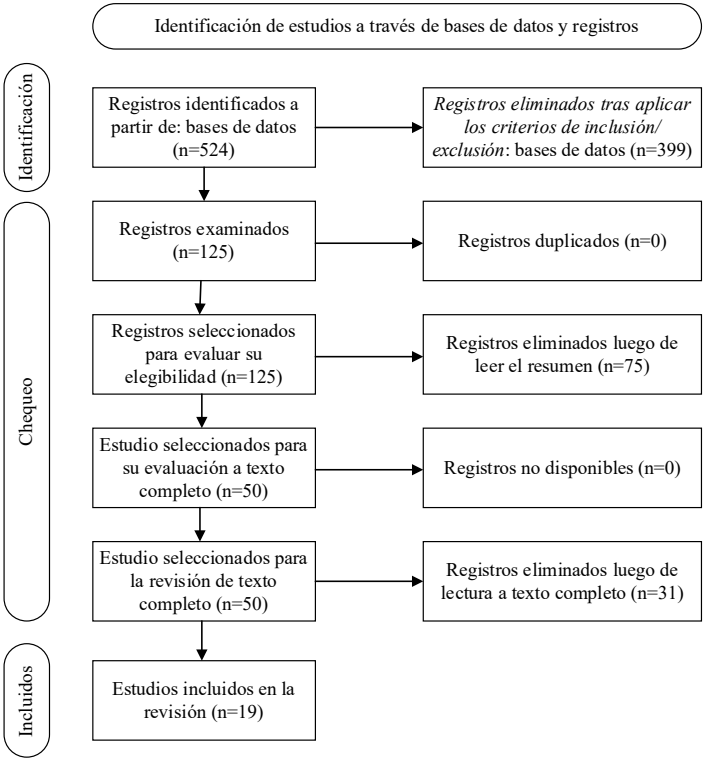
Cadena de búsqueda (03/11/2025)	(TITLE-ABS-KEY ("generative AI" OR "large language model*" OR LLM OR ChatGPT) AND TITLE-ABS-KEY (LLM OR "large language model*" OR ChatGPT) AND TITLE-ABS-KEY ("research lifecycle" OR "scholarly publishing" OR "research integrity" OR "peer review" OR authorship OR disclosure OR "scientific writing") AND TITLE-ABS-KEY (policy OR guideline OR "editorial policy" OR "author guidelines" OR governance OR transparency OR PRISMA OR "best practice*")) AND PUBYEAR > 2024 AND PUBYEAR < 2025 AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (LANGUAGE , "English")) AND (LIMIT-TO (OA , "all")) AND (LIMIT-TO (PUBSTAGE , "final"))

Etapa 2. Selección

La consulta recuperó 79 registro. Se realizó un cribado por título y resumen para pertinencia temática y, posteriormente, lectura a texto completo de los candidatos elegibles. La selección aplicó de forma consistente los criterios predefinidos y resolvió discrepancias por consenso. El conjunto final quedó constituido por 15 artículos que abordaban, con distinto énfasis, políticas editoriales, estudios empíricos y meta investigación, guías de reporte y propuestas metodológicas relacionadas con IA generativa en la comunicación científica (ver Figura 1, Tabla 2).

Figura 1.

Diagrama de flujo PRISMA que representa gráficamente el proceso de selección.



Etapa 3. Extracción de datos

Se desarrolló una plantilla estandarizada para registrar metadatos (título, revista, año, ámbito), tipo de contribución (política/posición, empírico/metainvestigación, guía/estándar, propuesta metodológica), etapa del ciclo impactada (búsqueda y síntesis, escritura, revisión por pares, reporte/comunicación, gobernanza editorial), requisitos de divulgación del uso de IA (herramienta, modelo/versión, propósito, prompts/semillas cuando procedía), postura sobre detectores y sus limitaciones, salvaguardas en revisión por pares (confidencialidad, supervisión humana, tareas permitidas y prohibiciones), estándares de reporte y exigencias mínimas, elementos de reproducibilidad y trazabilidad (control de versiones y auditoría de cambios), riesgos metodológicos identificados y recomendaciones operativas. La extracción fue pilotada en un subconjunto y verificada por un segundo revisor para asegurar consistencia. Dado el propósito exploratorio, no se aplicó una herramienta formal de riesgo de sesgo; sin embargo, se contextualizó la fuerza de la evidencia según la naturaleza de cada fuente y sus limitaciones declaradas.

Etapa 4. Síntesis

Se efectuó una síntesis narrativa temática, organizando los hallazgos en ejes previamente definidos y refinados inductivamente: gobernanza y divulgación del uso de IA, IA en revisión por pares, detección y su rol acotado, estándares de reporte y listas de verificación, y reproducibilidad con trazabilidad editorial. A partir de estos ejes se construyó un mapa etapa, riesgo, salvaguarda y se derivó una arquitectura mínima de adopción responsable con controles verificables orientados a universidades y revistas.

Resultados

La revisión integró quince estudios publicados entre 2024 y 2025, todos de acceso abierto, revisados por pares y centrados en el uso de inteligencia artificial generativa (IAg) o modelos de lenguaje de gran tamaño (LLM) en procesos de investigación y comunicación científica. Los artículos provienen de campos diversos, medicina, ciencias sociales, educación, ética, odontología y edición científica, lo que permite una visión transversal de los retos y convergencias que caracterizan la adopción responsable de la IA en el ciclo de la investigación. A partir del análisis temático se identificaron cuatro ejes principales: (1) gobernanza y divulgación del uso de IA; (2) IA en la revisión por pares; (3) detección y limitaciones de los verificadores automáticos; y (4) estándares de reporte y trazabilidad en la comunicación científica.

Gobernanza y divulgación del uso de IA

La mayoría de los estudios enfatiza la necesidad de establecer políticas editoriales claras que definan el alcance, los límites y la responsabilidad humana en el uso de IA durante la escritura y revisión científica. Yoo (2025) y Vasconcelos y Marušić (2025) coinciden en que la IA no puede figurar como autora, y su empleo debe revelarse explícitamente indicando la herramienta, modelo, versión y propósito. Se observa una transición desde declaraciones genéricas hacia esquemas de uso escalonado, en los que la responsabilidad recae en el autor principal y las instituciones patrocinadoras. Además, Lendvai (2025) muestra un aumento exponencial de publicaciones que incorporan ChatGPT o LLM, pero con reportes dispares de transparencia, lo que evidencia una falta de normalización en los formatos de divulgación. Los estudios proponen implementar matrices de revelación y registros auditables de prompts y versiones de modelo como mecanismos de trazabilidad y rendición de cuentas.

IA en la revisión por pares

Cinco artículos (Eggmann et al., 2025; Ebadi et al., 2025; Marrella et al., 2025; Mollaki, 2024; Lee et al., 2025) abordan directamente la integración de LLM en el proceso de revisión. Las evidencias convergen en que el apoyo de IA puede mejorar la consistencia lingüística y la detección de errores formales, pero conlleva riesgos de confidencialidad, sesgo y dependencia cognitiva. Eggmann et al. (2025) identifican el uso no declarado de IA en revisores de revistas odontológicas, mientras que Ebadi et al. (2025) evidencian posturas divididas: los revisores valoran la eficiencia pero temen comprometer la imparcialidad y la originalidad de los juicios. Marrella et al. (2025) comparan evaluaciones humanas y generadas por IA, concluyendo que los modelos pueden complementar, pero nunca sustituir, la evaluación experta. En conjunto, los estudios recomiendan protocolos con supervisión humana obligatoria, restricciones de carga de manuscritos inéditos a servicios públicos y delimitación precisa de tareas permitidas.

Detección y limitaciones de los verificadores automáticos

Erol et al. (2025) demuestran que los detectores de texto generados por IA presentan una alta tasa de falsos positivos y resultados inconsistentes entre plataformas, por lo que su uso como evidencia única en decisiones editoriales es inapropiado. La confiabilidad de estas herramientas depende del idioma, la longitud del texto y la versión del modelo base, factores que aún carecen de estandarización. En consecuencia, varios artículos (Yoo 2025; Teixeira da Silva, 2024; Mollaki, 2024) abogan por limitar los detectores a funciones de apoyo técnico y no sancionador, complementándolos con auditorías humanas y revisión cruzada.

Estándares de reporte y trazabilidad

Los trabajos de Huo et al. (2025) y Al-Qudimat et al. (2025) introducen avances en guías de reporte específicas, como el CHART Statement, que exige detallar herramienta, versión, propósito, configuración y supervisión humana. Estas iniciativas convergen con la propuesta de Yoo (2025) de incluir secciones obligatorias de "AI Disclosure" en manuscritos y con la sugerencia de Teixeira da Silva (2024) de adoptar listas de verificación universales equivalentes a PRISMA o CONSORT. La revisión revela la emergencia de estándares interdisciplinares de transparencia, aunque aún fragmentados por área y editorial. Paralelamente, estudios como los de Hu et al. (2025) y Calderón y Herrera (2025) advierten sobre riesgos metodológicos derivados de la inestabilidad de respuestas ante la repetición de prompts y sobre la necesidad de reproducibilidad mediante flujos de trabajo con control de versiones y auditoría de cambios.

Los hallazgos muestran que la integración responsable de la IA en el ciclo de la investigación avanza hacia un modelo de rendición de cuentas compartida. La tendencia más consolidada es la de exigir divulgación detallada y trazabilidad, seguida de la institucionalización de protocolos de revisión ética y de publicación transparente. Persisten, sin embargo, brechas notables: ausencia de métricas comparables de impacto, desigualdad de recursos entre regiones y disciplinas, y falta de interoperabilidad entre sistemas de reporte. La evidencia disponible sugiere que el desarrollo de una arquitectura mínima de adopción responsable, basada en divulgación verificable, responsabilidad humana, control editorial y trazabilidad reproducible, constituye la vía más viable para equilibrar innovación y confianza en la comunicación científica.

Discusión

Esta revisión de alcance evidencia una convergencia sustantiva en torno a cómo integrar la IA generativa en el ciclo de la investigación sin erosionar la integridad científica. Los estudios sintetizados, aunque heterogéneos en disciplinas y métodos, coinciden en un principio rector: la IA aporta eficiencia y claridad en tareas de apoyo, sobre todo en redacción, edición lingüística y organización de contenidos, pero su valor depende de una arquitectura de gobernanza que asegure divulgación granular, supervisión humana indelegable, confidencialidad en procesos sensibles y trazabilidad reproducible. La literatura reciente ya no discute si emplear IA, sino bajo qué condiciones, en qué etapas, y con qué garantías de transparencia y responsabilidad.

En primer lugar, se consolida una noción de gobernanza práctica basada en la revelación detallada del uso de IA. Se amplía el estándar de las meras "declaraciones de uso" hacia esquemas más finos que identifican herramienta, versión, fecha de uso, alcance de la tarea (por ejemplo, revisión de estilo frente a generación de texto original), parámetros relevantes y verificación humana. Este desplazamiento responde a dos tensiones: la no estacionariedad de los modelos, que obliga a datar y especificar versiones, y la necesidad de reproducibilidad, que requiere documentar prompts, plantillas o procedimientos cuando son parte sustantiva del método. En paralelo, la literatura normativa en edición científica converge en excluir la autoría de la IA y en asignar la responsabilidad última a los autores humanos y a las instituciones, reforzando un marco de rendición de cuentas que privilegia la trazabilidad por encima de la mera declaración retórica.

El segundo eje es la revisión por pares asistida por IA, donde las posturas son más matizadas. Se reconoce que los LLM pueden mejorar la consistencia lingüística y la claridad de las observaciones, especialmente cuando los revisores no son hablantes nativos del idioma de publicación. Sin embargo, se advierte que el uso sin control puede comprometer la confidencialidad de manuscritos inéditos, amplificar sesgos o inducir una dependencia cognitiva que empobrezca el juicio experto. De ahí que emerjan pautas de uso estrictamente instrumental: apoyo lingüístico y estructural sí; evaluación sustantiva, interpretación de resultados o recomendación editorial, no. Además, se subraya la importancia de entornos seguros (instancias locales, acuerdos de procesamiento de datos, restricciones contractuales) cuando el texto a tratar no es público, y se recomienda que los revisores declaren si utilizaron IA y con qué alcance, manteniendo la responsabilidad total de lo entregado.

Un tercer hallazgo transversal atañe a los detectores de texto IA. La evidencia sugiere que, pese a su utilidad exploratoria, presentan tasas de falsos positivos y variabilidad entre herramientas que impiden fundamentar decisiones sancionadoras solo en su dictamen. Factores como idioma, longitud de los textos y versiones de los modelos influyen de manera opaca en su rendimiento. La recomendación que se perfila es inequívoca: los detectores deben operar, en todo caso, como indicadores auxiliares, integrados en un enfoque probatorio múltiple que priorice la lectura crítica, la solicitud de materiales y la verificación independiente de fuentes y datos. Esta posición protege tanto a autores como a editores de decisiones precipitadas y reduce el riesgo de injusticias editoriales.

En cuarto lugar, se observa una emergencia de estándares de reporte específicos para el uso de IA, análogos a marcos consolidados como PRISMA o CONSORT, pero orientados a la transparencia algorítmica. Listas de verificación recientes exigen explicitar herramienta, versión, configuración, propósito, estrategia de prompts, corpus de referencia y naturaleza de la supervisión humana, además de ofrecer rutas para el

acceso a materiales, plantillas o registros de cambios cuando no existan restricciones legales o de privacidad. Esta tendencia sugiere que la comunidad está transitando de principios generales a implementaciones verificables, ampliando la caja de herramientas para auditar y comprender el aporte real de la IA al manuscrito o al proceso editorial.

Las implicaciones prácticas de esta síntesis son claras. Para autores, la adopción de una matriz de divulgación con control de versiones y verificación exhaustiva de referencias, DOIs y datos, no solo reduce riesgos, sino que agiliza el diálogo con editores y revisores. Para revisores y editores, la delimitación de usos permitidos y prohibidos de LLM, junto con políticas sobre confidencialidad y registro del proceso, habilita beneficios lingüísticos sin ceder en garantías éticas. Para CRAI y universidades, se vuelve estratégico ofrecer infraestructura segura (LLM locales o con acuerdos robustos), formación en "IA académica responsable" y plantillas interoperables de disclosure que armonicen manuscritos, repositorios y plataformas editoriales. Para financiadores y sociedades científicas, resulta oportuno incorporar secciones de IA en planes de gestión de datos y promover estudios comparativos que midan el impacto real de estas prácticas en tiempos editoriales, calidad de reporte y equidad de acceso.

Como toda síntesis, esta revisión tiene limitaciones. La evidencia disponible es reciente y, en buena medida, proviene de análisis de políticas, revisiones narrativas y estudios descriptivos, más que de ensayos controlados o auditorías ciegas a gran escala. La heterogeneidad conceptual, qué se entiende por "uso de IA", qué se considera "divulgación suficiente", dificulta la comparación directa. Y aunque el recorte temporal y de fuente (Scopus, acceso abierto, versión final, inglés) favorece actualidad y verificabilidad, puede sesgar contra experiencias valiosas en otros idiomas, contextos editoriales o regímenes de acceso. Estas restricciones, sin embargo, reflejan el estadio emergente del campo y justifican el enfoque de mapeo propio de una revisión de alcance.

A partir de aquí, se perfila una agenda de investigación concretable: desarrollar métricas comparables para evaluar la contribución de la IA en claridad, precisión factual y calidad de citación; realizar ensayos controlados de flujos con y sin IA, abarcando distintas disciplinas e idiomas; estudiar riesgos de confidencialidad en entornos cloud frente a on-prem y establecer estándares mínimos de seguridad y registro; analizar impactos distributivos de la adopción (brechas de infraestructura y licencias, especialmente en América Latina y otras regiones del Sur Global); y estandarizar formatos de prompts auditables (modelo, versión, parámetros y fecha) como parte de los materiales suplementarios cuando el método lo requiera.

Los hallazgos apuntan hacia un pacto operativo entre innovación y salvaguardas. La tríada de divulgación verificable, responsabilidad humana indelegable y trazabilidad reproducible, complementada con confidencialidad garantizada en revisión por pares y estándares de reporte específicos, constituye hoy la estrategia más plausible para capitalizar los beneficios de la IA sin menoscabar la integridad del registro científico. El desafío inmediato ya no es normativo, sino implementacional: convertir principios en procedimientos auditables, acompañados de métricas y evaluaciones independientes que permitan a revistas, CRAI e instituciones ajustar sus políticas sobre evidencia empírica y no solo sobre intuición o temor tecnológico (Mayorga, 2024). Solo así la IA dejará de ser un "riesgo a gestionar" para convertirse en un recurso confiable, alineado con los fines epistémicos y públicos de la ciencia.

Conclusiones

Esta revisión de alcance muestra que la integración de la IA generativa en el ciclo de la investigación ya no es una posibilidad periférica sino una realidad transversal (ver Tabla 3). La evidencia reciente converge en un principio rector: la IA puede aportar eficiencia, claridad y estandarización en tareas de apoyo (búsqueda, redacción, edición, estructuración y verificación), siempre que su uso se rija por divulgación verificable, responsabilidad humana indelegable, confidencialidad en procesos sensibles y trazabilidad reproducible. Sin estos pilares, el riesgo de opacidad, sesgo y erosión de la confianza supera los beneficios.

En el terreno editorial, las posiciones más sólidas descartan la autoría de la IA y recomiendan matrices de disclosure que documenten herramienta, versión, fecha, propósito y alcance del uso, así como la validación humana. En la revisión por pares, el consenso apunta a una utilización estrictamente instrumental (apoyo lingüístico/estilístico) y a la prohibición de cargar manuscritos inéditos en servicios públicos, privilegiando entornos locales o con garantías contractuales. Respecto a los detectores de texto IA, la recomendación es inequívoca: se aceptan como indicadores auxiliares, pero no como prueba suficiente para decisiones sancionadoras.

Emergen, además, estándares de reporte específicos (análogos a PRISMA o CONSORT) que incorporan la transparencia algorítmica: identificación del modelo y versión, estrategia de prompts cuando sea parte sustantiva del método, supervisión humana y disponibilidad de materiales. Esta transición de principios a implementaciones auditables es la señal más clara de madurez del campo y el camino más directo para alinear innovación con integridad científica.

Para la práctica, proponemos un mínimo operacional aplicable entre disciplinas: (1) disclosure granular y fechado de todo uso de IA; (2) control de versiones y registro de cambios en manuscritos y revisiones; (3) verificación humana obligatoria de hechos, citas y DOIs; (4) políticas institucionales y editoriales que delimiten usos permitidos/prohibidos en revisión por pares y aseguren confidencialidad; y (5) adopción progresiva de checklists de reporte de IA integrables en plataformas editoriales y repositorios.

Este trabajo presenta limitaciones propias de su naturaleza y del estado emergente de la evidencia: predominio de estudios descriptivos/políticos sobre ensayos controlados, heterogeneidad en definiciones de “uso de IA” y focalización en literatura reciente, de acceso abierto y en inglés. Aun así, el mapeo ofrece una base robusta para decisiones editoriales, institucionales y de política científica en el corto plazo.

Como agenda inmediata, se priorizan tres frentes: métricas comparables del impacto real de la IA en calidad y tiempos editoriales; ensayos controlados de flujos con/sin IA que incluyan distintos idiomas y áreas; y estándares mínimos de seguridad y auditoría para revisión por pares y manejo de manuscritos no públicos. Avanzar en estos ejes permitirá pasar de un debate normativo a una implementación verificable, sosteniendo la promesa de la IA como recurso confiable al servicio de los fines epistémicos y públicos de la ciencia.

Referencias Bibliográficas

Akinrinola, O., Okoye, C. C., Ofodile, O. C., & Ugochukwu, C. E. (2024). Navigating and reviewing ethical dilemmas in AI development: Strategies for transparency, fairness, and accountability. GSC Advanced Research and Reviews, 18(3), 050-058. https://doi.org/10.30574/gscarr.2024.18.3.0088

Alkhawam, M., Almobayed, A., Pandey, A., Ebrahimi, A. J., & Ahmed, M. I. (2025). Exploring AI use policies in manuscript writing in cardiology and vascular journals. *American Heart Journal Plus: Cardiology Research and Practice, 58*, 100586. https://doi.org/10.1016/j.ahjo.2025.100586

Al-Qudimat, A. R., Fares, Z. E., Elaarag, M., Al-Zoubi, R. M., & Aboumarzouk, O. M. (2025). Advancing medical research through artificial intelligence: Progressive and transformative strategies A literature review. *Health Science Reports, 8*(2), e70200. https://doi.org/10.1002/hsr2.70200

Al-Moteri, M. (2025). Scenario-based forecasting of ChatGPT's role as a research assistant in nursing by 2030. *International Nursing Review, 72*(3), e70055. https://doi.org/10.1111/inr.70055

Calderon, R., & Herrera, F. (2025). And Plato met ChatGPT: An ethical reflection on the use of chatbots in scientific research writing, with a particular focus on the social sciences. *Humanities and Social Sciences Communications, 12*(1), 713. https://doi.org/10.1057/s41599-025-04650-0

Carter, R. E., Attia, Z. I., Lopez-Jimenez, F., & Friedman, P. A. (2019). Pragmatic considerations for fostering reproducible research in artificial intelligence. NPJ digital medicine, 2(1), 42. https://doi.org/10.1038/s41746-019-0120-2

Echeverría, C. A., Flores, M. I., Vaquerano, J. A., de Salazar, K. Y. R., Maravilla, J. F. P., García-de-Ridruejo, J. G., & París, F. X. G. (2023). Los factores de riesgos psicosociales relacionales en las personas trabajadoras de las empresas salvadoreñas. Interconectando Saberes, (16), 151-163. https://doi.org/10.25009/is.v0i16.2811

Ebadi, S., Nejadghanbar, H., Salman, A. R., & Khosravi, H. (2025). Exploring the impact of generative AI on peer review: Insights from journal reviewers. *Journal of Academic Ethics, 23*(3), 1383–1397. https://doi.org/10.1007/s10805-025-09604-4

Eggmann, F., Hildebrand, H., & Bornstein, M. M. (2025). Who reviewed this? Toward responsible integration of large language models for peer review of scientific articles in dental medicine. *Swiss Dental Journal, 135*(3), 1–15. https://doi.org/10.61872/sdj-2025-03-01

Erol, G., Ergen, A., Gülşen Erol, B., Sevgi, U. T., & Güngör, A. (2025). Can we trust academic AI detectors? Accuracy and limitations of AI-output detectors.

Acta Neurochirurgica, 167(1), 214. https://doi.org/10.1007/s00701-025-06622-4

Hernández, R. M. F., Mancía, G. A. P., Mayorga, C. A. E., & Mendoza, N. O. O. (2023). La evaluación de los aprendizajes de las carreras de computación y afines, en la educación universitaria salvadoreña. Revista Integración, (11), 25-50. https://doi.org/10.5377/ri.v1i11.18377

Huo, B., Collins, G., Chartash, D., Guallar, E., & Guyatt, G. (2025). Reporting guideline for chatbot health advice studies: The CHART statement. *JAMA Network Open*. https://doi.org/10.1001/jamanetworkopen.2025.30220

Hu, H., Zhou, Q., & Hashim, H. (2025). Negotiating identity in the age of ChatGPT: Non-native English researchers' experiences with AI-assisted academic writing. *Humanities and Social Sciences Communications, 12*(1), 965. https://doi.org/10.1057/s41599-025-05351-4

Istasy, P., Lee, W. S., Iansavitchene, A., Upshur, R., Sadikovic, B., Lazo-Langner, A., & Chin-Yee, B. (2021). The impact of artificial intelligence on health equity in oncology: a scoping review. Blood, 138, 4934. https://doi.org/10.1182/blood-2021-149264

Lee, J., Lee, J., & Yoo, J.-J. (2025). The role of large language models in the peer-review process: Opportunities and challenges for medical journal reviewers and editors. *Journal of Educational Evaluation for Health Professions, 22*, 4. https://doi.org/10.3352/jeehp.2025.22.4

Lendvai, G. F. (2025). ChatGPT in academic writing: A scientometric analysis of literature published between 2022 and 2023. *Journal of Empirical Research on Human Research Ethics, 20*(3), 131–148. https://doi.org/10.1177/15562646251350203

Li, B., Qi, P., Liu, B., Di, S., Liu, J., Pei, J., ... & Zhou, B. (2023). Trustworthy AI: From principles to practices. ACM Computing Surveys, 55(9), 1-46. https://doi.org/10.1145/3555803

Lu, Q., Zhu, L., Xu, X., Whittle, J., Zowghi, D., & Jacquet, A. (2024). Responsible AI pattern catalogue: A collection of best practices for AI governance and engineering. ACM Computing Surveys, 56(7), 1-35. https://doi.org/10.1145/3626234

Marrella, D., Jiang, S., Ipaktchi, K., & Liverneaux, P. (2025). Comparing AI-generated and human peer reviews: A study on 11 articles. *Hand Surgery and Rehabilitation, 44*(4), 102225. https://doi.org/10.1016/j.hansur.2025.102225

Mayorga, C. A. E. (2024). Teaching Attitudes and the Participation of Individuals with Physical Motor Disabilities in Higher Education in El Salvador. Journal of Multidisciplinary Studies in Human Rights & Science (JMSHRS), (6). https://doi.org/10.5281/zenodo.13384325

Mayorga, C. A. E., de Guardado, K. A. E., & Mendoza, N. O. O. (2023). Alfabetización tecnológica del docente universitario en El Salvador. Revista de Investigación, 2(14), 10-29. https://doi.org/10.5377/revunivo.v2i14.17012

Mollaki, V. (2024). Death of a reviewer or death of peer review integrity? The challenges of using AI tools in peer reviewing and the need to go beyond publishing policies. *Research Ethics, 20*(2), 239–250. https://doi.org/10.1177/17470161231224552

Polanco, M. I. F., & Mayorga, C. A. E. (2025). Scientific Production in Central America (1996–2023): Bibliometric Analysis of Regional Trends, Collaboration, and Research Impact. Publications, 13(3), 44. https://doi.org/10.3390/publications13030044

Teixeira da Silva, J. A. (2024). Proposing authorship for artificial intelligence and large language models. *European Science Editing, 50*, e123908. https://doi.org/10.3897/ese.2024.e123908

Teodoro, D., Naderi, N., Yazdani, A., Zhang, B., & Bornet, A. (2025). A scoping review of artificial intelligence applications in clinical trial risk assessment. npj Digital Medicine, 8(1), 486. https://doi.org/10.1038/s41746-025-01886-7

Yoo, J.-H. (2025). Defining the boundaries of AI use in scientific writing: A comparative review of editorial policies. *Journal of Korean Medical Science, 40*(23), e187. https://doi.org/10.3346/jkms.2025.40.e187

ANEXOS

Anexo 1

Tabla 2.

Estudios incluidos en la revisión de alcance (2024–2025)

N°	Autor(es) y año	Revista / Área	Tipo de contribución	Etapa del ciclo de investigación	Principales hallazgos / recomendaciones
1	Yoo (2025)	Journal of Korean Medical Science / Medicina	Revisión de políticas editoriales	Gobernanza, reporte	Propone lineamientos flexibles pero firmes sobre divulgación y responsabilidad humana; destaca la insuficiencia de detectores automáticos.
2	Huo et al. (2025)	British Journal of Surgery /	Guía de reporte (CHART statement)	Reporte y transparencia	Desarrolla una lista de verificación (CHART) para reportar estudios

N°	Autor(es) y año	Revista / Área	Tipo de contribución	Etapa del ciclo de investigación	Principales hallazgos / recomendaciones
		Ciencias de la salud			de IA conversacional; promueve transparencia y trazabilidad metodológica.
3	Eggmann et al. (2025)	Swiss Dental Journal / Odontología	Estudio transversal de políticas de revisión	Revisión por pares	Identifica vacíos en el uso de IA en revisión dental; recomienda guías de confidencialidad y responsabilidad del revisor.
4	Erol et al. (2025)	Acta Neurochirurgica / Neurociencias	Evaluación empírica	Integridad, detección	Analiza fiabilidad de detectores de texto IA; concluye que presentan alta tasa de falsos positivos y deben usarse con cautela.
5	Calderón & Herrera (2025)	Humanities and Social Sciences Communications / Ciencias sociales	Reflexión ética interdisciplinaria	Escritura científica	Cuestiona la autoría híbrida humano-IA y defiende la necesidad de marcos éticos multidisciplinares.
6	Hu et al. (2025)	Humanities and Social Sciences Communications / Educación e identidad	Estudio cualitativo	Escritura científica	Explora tensiones identitarias en investigadores no nativos; recomienda alfabetización ética en IA académica.
7	Ebadi et al. (2025)	Journal of Academic Ethics / Ética académica	Estudio cualitativo	Revisión por pares	Identifica beneficios y dilemas del uso de LLMs en revisión; enfatiza la supervisión humana y guías claras.

N°	Autor(es) y año	Revista / Área	Tipo de contribución	Etapa del ciclo de investigación	Principales hallazgos / recomendaciones
8	Marrella et al. (2025)	Hand Surgery and Rehabilitation / Medicina	Estudio comparativo	Revisión por pares	Compara revisiones humanas y generadas por IA; concluye que IA puede complementar revisiones con control experto.
9	Lendvai (2025)	Journal of Empirical Research on Human Research Ethics / Ética en investigación	Análisis bibliométrico y temático	Gobernanza, reporte	Mapea la estructura temática sobre ChatGPT en escritura académica; identifica necesidad de políticas globales de transparencia.
10	Vasconcelos & Marušić (2025)	EMBO Reports / Ciencia y ética	Ensayo crítico	Integridad y políticas	Discute la redefinición de integridad científica ante IA generativa; urge actualizar definiciones normativas.
11	Lee et al. (2025)	Journal of Educational Evaluation for Health Professions / Educación médica	Revisión teórica	Revisión por pares	Analiza potencial de LLMs para asistencia en revisión médica; recomienda complementariedad bajo guías éticas.
12	Al-Moteri (2025)	International Nursing Review / Enfermería	Prospectiva de escenarios	Gobernanza e investigación	Plantea escenarios éticos y colaborativos para integración responsable de IA en investigación en salud hacia 2030.
13	Al-Qudimat et al. (2025)	Health Science Reports / Medicina	Revisión de literatura	Escritura y publicación	Explora impacto de IA en redacción médica; resalta beneficios de eficiencia y desafíos

N°	Autor(es) y año	Revista / Área	Tipo de contribución	Etapa del ciclo de investigación	Principales hallazgos / recomendaciones
					éticos de autoría y sesgo.
14	Teixeira da Silva (2024)	European Science Editing / Edición científica	Ensayo conceptual	Autoría y gobernanza	Propone debate sobre la posibilidad de 'autoría IA'; argumenta apertura responsable dentro de límites éticos.
15	Mollaki (2024)	Research Ethics / Ética en publicación	Revisión crítica	Revisión por pares	Advierte falta de políticas sobre IA en evaluación; propone protocolos transparentes más allá de políticas genéricas.

Tabla 3.

Arquitectura mínima de adopción responsable de IA en el ciclo de la investigación

Etapa del ciclo de investigación	Riesgos principales	Salvaguardas y controles propuestos	Evidencia o fuente clave
Búsqueda y síntesis bibliográfica	Referencias falsas o desactualizadas; dependencia acrítica de resúmenes generados por IA; pérdida de trazabilidad de fuentes.	Uso de motores con recuperación aumentada (RAG) y referencias verificables; registro de prompts; versión y fecha del modelo; validación cruzada de fuentes.	Yoo (2025); Lendvai (2025); Vasconcelos & Marušić (2025).
Redacción y edición de manuscritos	Plagio inadvertido; autoría ambigua; citaciones incorrectas; opacidad sobre el rol de la IA.	Disclosure granular (herramienta, versión, fecha, propósito, alcance); control de versiones y registro de cambios; validación humana de DOIs, datos y citas.	Teixeira da Silva (2024); Calderón & Herrera (2025); Hu et al. (2025).
Revisión por pares	Riesgo de fuga de información; sesgos o juicios generados por IA; pérdida de responsabilidad humana.	Uso acotado a apoyo lingüístico/estructural; prohibición de cargar manuscritos inéditos en servicios públicos; declaración de uso por revisores; revisión con supervisión humana.	Eggmann et al. (2025); Ebadi et al. (2025); Marrella et al. (2025); Mollaki (2024).

Reporte y comunicación de resultados	Inconsistencias en transparencia; ausencia de identificación del modelo o versión usada; reproducibilidad limitada.	Aplicación de listas de verificación específicas (p. ej., CHART); mención explícita de modelo, versión, prompts y validación humana; anexos con registro de cambios o seeds.	Huo et al. (2025); Al-Qudimat et al. (2025); Yoo (2025).
Gobernanza editorial y políticas	Uso desigual de detectores; falta de armonización entre revistas; sanciones basadas en herramientas no validadas.	Prohibir uso sancionador de detectores; aplicar revisión humana complementaria; publicar políticas institucionales y matrices de divulgación obligatorias.	Erol et al. (2025); Vasconcelos & Marušić (2025); Mollaki (2024).

Nota: resume los principales riesgos identificados en cada etapa del ciclo de la investigación, junto con las salvaguardas y controles verificables propuestos para una adopción responsable de IA generativa según la evidencia revisada (2024–2025).