

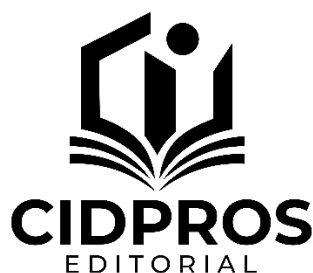
NUEVAS FORMAS DE BUSCAR Y ANALIZAR CON IA LA INFORMACIÓN CIENTÍFICA

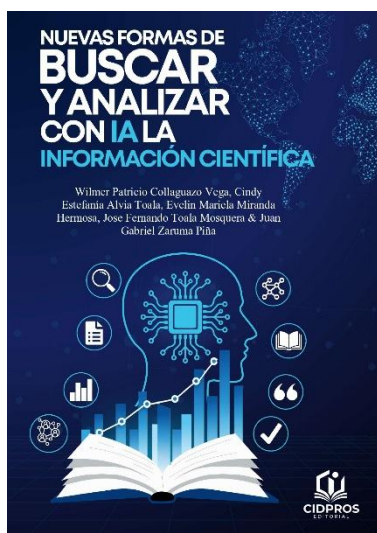
Wilmer Patricio Collaguazo Vega, Cindy
Estefanía Alvia Toala, Evelin Mariela Miranda
Hermosa, Jose Fernando Toala Mosquera & Juan
Gabriel Zaruma Piña



Nuevas formas de buscar y analizar con IA la información científica

Wilmer Patricio Collaguazo Vega, Cindy Estefanía
Alvia Toala, Evelin Mariela Miranda Hermosa, Jose
Fernando Toala Mosquera & Juan Gabriel Zaruma Piña





Datos bibliográficos

ISBN	978-9907-9574-8-8
Título del libro	Nuevas formas de buscar y analizar con IA la información científica
Autores	Wilmer Patricio Collaguazo Vega Cindy Estefanía Alvia Toala Evelin Mariela Miranda Hermosa Jose Fernando Toala Mosquera Juan Gabriel Zaruma Piña
Editorial	CIDPROS EDITORIAL
Materia	001.4 - Investigación
Público objetivo	Profesional / académico
Año	2026
Número de edición	1
Tamaño	2,74Mb
Soporte	Libro digital descargable
Formato	PDF (.pdf)
Idioma	Español
DOI	https://doi.org/10.67166/34zqjv63

Hecho en Ecuador / Made in Ecuador

MBA. Wilmer Patricio Collaguazo Vega

Universidad Tecnológica Indoamérica, Facultad de Ciencias de la Educación

wilmercollaguazo@uti.edu.ec

<https://orcid.org/0000-0001-6962-9171>

Latacunga, Cotopaxi, Ecuador

Semblanza

Wilmer Patricio Collaguazo Vega es docente universitario radicado en la ciudad de Latacunga, Ecuador. Es Magíster en Docencia Universitaria y Pedagogía del Inglés como Lengua Extranjera y Máster en Alta Dirección. Actualmente complementa su formación con estudios doctorales en el ámbito de la educación y la innovación, orientados al fortalecimiento de su perfil académico y profesional.

Cuenta con más de diez años de experiencia en educación superior y se desempeña como docente de grado y posgrado. Su trayectoria incluye la enseñanza del inglés, la formación de futuros profesionales de la educación, el diseño curricular, la evaluación de los aprendizajes, el acompañamiento académico y la gestión educativa. En la Universidad Tecnológica Indoamérica ha asumido diversas responsabilidades vinculadas con el desarrollo académico y el mejoramiento de la calidad educativa.

Entre sus principales intereses se encuentra la aplicación de la inteligencia artificial en la educación, con especial énfasis en la ingeniería de prompts y su utilidad para el diseño de experiencias de aprendizaje, materiales educativos y actividades académicas. Su labor se sustenta en la actualización permanente, la innovación pedagógica y el compromiso con una educación inclusiva y pertinente, procurando que las tecnologías emergentes se integren de manera responsable y con un propósito educativo claro.



Ing. Cindy Estefanía Alvia Toala

Investigador Independiente

estefyalto@gmail.com

<https://orcid.org/0009-0004-0123-2142>

Manta, Manabí, Ecuador

Semblanza

Cindy Estefanía Alvia Toala, ingeniera en Comercio Exterior y Negocios Internacionales, es originaria de Manta, Manabí, donde actualmente reside. Posee una sólida formación académica, destacándose como Magíster en Ciencias Sociales con mención en Investigación para el Desarrollo Social por la Universidad Laica Eloy Alfaro de Manabí. Su trayectoria profesional se centra en el ámbito educativo y de investigación social, desempeñándose como Técnico Docente en la Universidad Laica Eloy Alfaro de Manabí, donde ha desarrollado e implementado planes de nivelación académica y estrategias de enseñanza personalizadas. Además, ha contribuido significativamente en la publicación de guías de estudio en áreas como Matemática Administrativa y Economía Global. Cindy es reconocida por su capacidad de liderazgo, innovación y compromiso con el desarrollo social sostenible. Su enfoque profesional está orientado a maximizar el potencial de sus estudiantes y a fomentar el bienestar ocupacional, reflejando un profundo compromiso con el impacto positivo en su comunidad.



Mgtr. Evelin Mariela Miranda Hermosa

Investigador Independiente

evemirandah@gmail.com

<https://orcid.org/0009-0008-1708-3945>

Quito, Pichincha, Ecuador

Semblanza

Evelin Mariela Miranda Hermosa, profesional ecuatoriana, es docente, comunicadora y especialista en publicidad y marketing digital. Es Magíster en Inteligencia Artificial para la Educación por la Universidad Estatal de Milagro y Magíster en Publicidad Digital por la Universidad Tecnológica Equinoccial; además, es Licenciada en Publicidad por la misma institución y Tecnóloga en Medios Masivos por la Universidad San Francisco de Quito.

Cuenta con una trayectoria profesional vinculada a la comunicación, publicidad, marketing digital y educación superior. Desde 2016 se desempeña como docente y, desde 2019, como Coordinadora de la Comisión de Titulación del Instituto Superior - INSTA, donde coordina procesos de titulación. También ha ejercido funciones de coordinación de vinculación con la sociedad y ha participado como docente en la Universidad Tecnológica Equinoccial, Universitario de Formación y en la Universidad Técnica del Norte, en el programa de Maestría en Marketing Digital.

Su trayectoria investigativa se orienta hacia la integración de la inteligencia artificial, las tecnologías digitales y las estrategias innovadoras en los procesos educativos. Es coautora de publicaciones científicas relacionadas con la inteligencia artificial y la inclusión educativa, el uso de redes sociales en el proceso de enseñanza-aprendizaje y la participación estudiantil. Su labor se caracteriza por el compromiso con la educación superior, la innovación y la incorporación responsable de las tecnologías emergentes, buscando contribuir al fortalecimiento de las prácticas docentes y al desarrollo de procesos educativos acordes con las transformaciones de la sociedad contemporánea.



Jose Fernando Toala Mosquera MD ESP MSC PHDe

Investigador Independiente

josetoalamosquera@gmail.com

<https://orcid.org/0000-0002-2188-0065>

Milagro, Guayas, Ecuador

Semblanza

José Fernando Toala Mosquera es médico ocupacional y docente universitario con 23 años de experiencia profesional. Su trayectoria se ha desarrollado principalmente en el campo de la medicina del trabajo, las urgencias y emergencias, la higiene industrial y los trabajos de alto riesgo ocupacional, integrando conocimientos orientados a la prevención, evaluación y atención de condiciones relacionadas con la salud y la seguridad en el ámbito laboral.

Su formación y experiencia también comprenden áreas como la ergonomía ocupacional y el diseño de puestos de trabajo, así como la psicología organizacional y del trabajo. Estas áreas han fortalecido su desempeño profesional en contextos donde resulta necesario analizar las condiciones de trabajo, los factores de riesgo y la organización de los espacios laborales desde una perspectiva integral.

Es egresado del PhD en docencia de la Universidad Nacional de Rosario, Argentina, y cuenta además con formación como auditor de sistemas integrados de gestión ISO y BPM. Ha realizado estudios de posgrado en Ecuador, Perú, Chile y Argentina, ampliando su preparación académica y profesional en distintas áreas vinculadas con su ejercicio.

Su experiencia como médico ocupacional y docente universitario le ha permitido articular la práctica profesional con la formación académica, manteniendo una trayectoria vinculada con la salud ocupacional, la prevención de riesgos, la gestión y la docencia universitaria.



Lcdo. Juan Gabriel Zaruma Piña

Investigador Independiente

juan.zaruma@educacion.gob.ec

<https://orcid.org/0009-0008-6936-7499>

Santiago, Morona Santiago, Ecuador

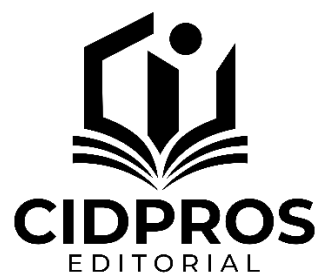
Semblanza

Juan Gabriel Zaruma Piña es un destacado educador ecuatoriano cuya trayectoria profesional y vocación de servicio están fuertemente enraizadas en la transformación pedagógica de la educación en zonas rurales. Su convicción humanista entiende el aula rural no solo como un espacio de instrucción formal, sino como un eje dinámico de equidad, inclusión social y desarrollo comunitario. Desde el inicio de su labor profesional, ha guiado su práctica pedagógica por una profunda sensibilidad hacia los retos del entorno escolar local, promoviendo el conocimiento integral como un instrumento clave para generar oportunidades reales de progreso social y bienestar colectivo.



Posee una sólida y diversa formación académica: es Licenciado en Ciencias de la Educación Básica (2do a 7mo grado) por la Universidad de Cuenca, trayectoria formativa que enriqueció con una experiencia de intercambio académico internacional durante el séptimo semestre en el Instituto Tecnológico de Sonora (ITSON) en México. Fortaleciendo de manera continua su perfil profesional, actualmente cursa la Maestría en Tecnología Educativa y Competencias Digitales en TECH Universidad Tecnológica. Esta preparación se complementa con un permanente proceso de actualización en didáctica STEAM, enfoques pedagógicos subnivel EGB, inclusión, atención a la diversidad y herramientas digitales especializadas como el programa Google Workspace for Education Plus.

En el ámbito laboral, ejerció la docencia en diversos niveles de la educación básica durante tres años, consolidando un profundo conocimiento de las dinámicas del aula y de la realidad estudiantil. Actualmente se desempeña como rector encargado de la Unidad Educativa Real Audiencia de Quito, donde ejerce un liderazgo institucional orientado a la excelencia académica, la gestión eficiente y el fortalecimiento continuo del trabajo docente. Su visión investigativa e intelectual se refleja en la publicación del artículo científico titulado “El bullying y las estrategias de mediación entre pares en instituciones de educación básica”, así como en la autoría del libro “Aprender a pensar en la era digital: Pensamiento crítico para niños y adolescentes”. Docente y directivo comprometido, encarna valores de ética, resiliencia e innovación, impulsando día a día la construcción de un sistema educativo más justo, inclusivo y vanguardista.



El contenido y las ideas expuestas en esta obra se encuentran protegidos por la normativa vigente en materia de propiedad intelectual y constituyen derechos exclusivos de su(s) autor(es).

Todos los derechos reservados © 2026

Sinopsis

Nuevas formas de buscar y analizar con IA la información científica examina cómo la inteligencia artificial está modificando las prácticas de búsqueda, organización, lectura, análisis y verificación de literatura académica. La obra aborda la evolución desde procedimientos basados principalmente en palabras clave hacia sistemas de recuperación semántica, recomendación automatizada y procesamiento contextual de documentos. También analiza la construcción de mapas digitales del conocimiento mediante redes de autores, documentos, conceptos, cocitación y agrupamientos temáticos, destacando la necesidad de interpretar críticamente estas representaciones. Otro eje corresponde a la lectura automatizada de grandes colecciones, incluyendo extracción de información, resumen, clasificación, comparación y síntesis de evidencia. El libro estudia además el uso de inteligencia artificial para identificar tendencias, temas emergentes, vacíos de conocimiento, patrones metodológicos y contradicciones científicas. Finalmente, desarrolla criterios para verificar fuentes, referencias, DOI y correspondencia entre afirmaciones y citas, considerando problemas como alucinaciones, errores bibliográficos y pérdida de trazabilidad. El enfoque mantiene una posición crítica: la automatización amplía la capacidad de procesar información científica, pero no sustituye la evaluación metodológica, la lectura contextual ni la responsabilidad del investigador sobre la evidencia utilizada.

Palabras clave: inteligencia artificial; información científica; búsqueda semántica; análisis bibliométrico; verificación de fuentes.

Synopsis

New Ways to Search and Analyze Scientific Information with AI examines how artificial intelligence is transforming the practices of searching, organizing, reading, analyzing, and verifying academic literature. The book addresses the transition from procedures based mainly on keywords to systems for semantic retrieval, automated recommendation, and contextual document processing. It also analyzes the construction of digital knowledge maps through networks of authors, documents, concepts, co-citation, and thematic clustering, emphasizing the need for critical interpretation of these representations. Another central focus is the automated reading of large collections, including information extraction, summarization, classification, comparison, and evidence synthesis. The book also explores the use of artificial intelligence to identify trends, emerging topics, knowledge gaps, methodological patterns, and scientific contradictions. Finally, it develops criteria for verifying sources, references, DOIs, and the correspondence between claims and citations, considering problems such as hallucinations, bibliographic errors, and loss of traceability. The approach maintains a critical stance: automation expands the capacity to process scientific information, but it does not replace methodological evaluation, contextual reading, or the researcher's responsibility for the evidence used.

Keywords: artificial intelligence; scientific information; semantic search; bibliometric analysis; source verification.

Índice de Contenido

Sinopsis	10
Synopsis	11
Introducción	14
Capítulo 1. Búsqueda avanzada de literatura científica con inteligencia artificial	17
1.1 Evolución de la búsqueda científica en entornos digitales.....	18
1.2 Formulación inteligente de preguntas y estrategias de búsqueda	20
1.3 Búsqueda semántica y recuperación contextual de información	23
1.4 Herramientas de inteligencia artificial para localizar literatura científica	27
1.5 Refinamiento, filtrado y priorización de resultados	29
1.6 Evaluación de pertinencia y calidad de los documentos recuperados.....	32
1.7 Diseño de estrategias reproducibles de búsqueda asistida por IA.....	35
Capítulo 2. Construcción de mapas digitales del conocimiento científico	39
2.1 Organización del conocimiento científico mediante representaciones digitales.....	40
2.2 Redes de autores, documentos y conceptos científicos.....	42
2.3 Análisis de coautoría, cocitación y acoplamiento bibliográfico	45
2.4 Identificación automática de conceptos y relaciones temáticas	48
2.5 Agrupamiento de documentos y formación de clústeres científicos.....	51
2.6 Visualización interactiva de campos y líneas de investigación.....	54
2.7 Interpretación crítica de mapas científicos generados con IA.....	57
Capítulo 3. Lectura automatizada de grandes colecciones de artículos	60
3.1 Procesamiento automatizado del contenido científico	61
3.2 Extracción de información relevante de textos académicos.....	63
3.3 Resumen automático de artículos y documentos científicos.....	66
3.4 Comparación simultánea de múltiples investigaciones.....	69
3.5 Clasificación temática y categorización automatizada.....	72
3.6 Síntesis de evidencia mediante inteligencia artificial	75
3.7 Riesgos y límites de la lectura automatizada de literatura científica	78

Capítulo 4. Detección de tendencias, vacíos y contradicciones científicas	81
4.1 Identificación de tendencias en la producción científica	82
4.2 Evolución temporal de temas y líneas de investigación.....	84
4.3 Detección de temas emergentes mediante inteligencia artificial	87
4.4 Identificación de vacíos de conocimiento	90
4.5 Reconocimiento de resultados divergentes y contradicciones científicas.....	92
4.6 Análisis de patrones metodológicos y conceptuales	95
4.7 Uso de tendencias y vacíos para formular nuevas investigaciones	98
Capítulo 5. Verificación inteligente de fuentes, citas y referencias.....	102
5.1 Calidad y confiabilidad de las fuentes científicas	103
5.2 Verificación de existencia y correspondencia de referencias.....	105
5.3 Comprobación automatizada de autores, títulos, revistas y DOI	108
5.4 Evaluación de la correspondencia entre afirmaciones y citas	111
5.5 Identificación de referencias falsas, incompletas o inconsistentes	113
5.6 Trazabilidad de la evidencia científica utilizada	116
5.7 Uso responsable de inteligencia artificial en la gestión de citas y referencias.....	119
Conclusión	122
Bibliografía	125

Introducción

La producción científica contemporánea se desarrolla en un entorno caracterizado por una disponibilidad documental que supera ampliamente las posibilidades de revisión manual. Bases de datos, repositorios, revistas digitales y sistemas de información concentran cantidades crecientes de publicaciones, mientras las herramientas de inteligencia artificial incorporan nuevas formas de localizar, relacionar y procesar esos contenidos. Esta transformación modifica una actividad fundamental de la investigación: pasar de disponer de información a identificar qué evidencia resulta pertinente, cómo se organiza y bajo qué condiciones puede utilizarse de manera científicamente justificable.

La búsqueda de literatura constituye uno de los ámbitos donde este cambio resulta más visible. Los procedimientos centrados exclusivamente en palabras clave conviven ahora con sistemas de recuperación semántica, recomendación académica y consultas formuladas mediante lenguaje natural. Gottardi et al. (2021) muestran que la búsqueda semántica amplía la encontrabilidad de información científica al considerar relaciones de significado, mientras Ajith et al. (2024) sitúan la recuperación de literatura como una tarea especializada que requiere identificar documentos pertinentes frente a necesidades científicas concretas. Estas posibilidades amplían la exploración, pero también exigen mayor precisión metodológica.

La inteligencia artificial interviene igualmente después de localizar los documentos. El cribado de títulos y resúmenes, la extracción de información, la clasificación temática y la síntesis automatizada permiten trabajar con colecciones que serían difíciles de procesar de manera exclusivamente manual. Ofori et al. (2024) examinan la automatización de distintas tareas vinculadas con las revisiones sistemáticas, mientras Tang et al. (2023) analizan la capacidad de los modelos de lenguaje para sintetizar evidencia científica. El desafío consiste en aprovechar esta capacidad sin eliminar el contexto metodológico del que dependen los resultados originales.

Otra transformación corresponde a la posibilidad de representar la literatura como una red de relaciones. Los documentos pueden vincularse mediante autoría, citas, conceptos, temas y otros atributos que permiten construir mapas digitales del conocimiento. Ávila (2024) muestra cómo los grafos pueden aplicarse a la visualización de datos abiertos de investigación, y Verma et al. (2023) estudian los grafos de conocimiento académico como estructuras para organizar la comunicación científica. Estas representaciones permiten observar

configuraciones amplias del campo, aunque su interpretación depende de la calidad y cobertura de los datos utilizados.

El análisis de grandes corpus también abre posibilidades para estudiar cómo cambia la ciencia. Los modelos temáticos pueden utilizarse para identificar agrupamientos, observar su evolución temporal y detectar conexiones entre áreas que antes aparecían separadas. Wu et al. (2024) examinan el desarrollo de modelos neuronales de tópicos, mientras Vázquez et al. (2022) muestran la importancia de validar las estructuras temáticas mediante información procedente de grafos y metadatos documentales. Estas herramientas permiten reconocer señales relevantes, pero una frecuencia elevada, un clúster nuevo o una zona poco poblada no constituyen por sí mismos una conclusión científica.

La identificación de vacíos y contradicciones requiere precisamente superar esa lectura automática de los patrones. Arias y Artigas (2022) vinculan las brechas de investigación con la formulación de problemas científicos relevantes, lo que exige determinar qué conocimiento falta realmente y por qué esa ausencia merece ser estudiada. De forma complementaria, Wadden et al. (2022) muestran que la verificación de afirmaciones científicas puede relacionar proposiciones con evidencia que las respalda o contradice. La inteligencia artificial puede ayudar a localizar estas relaciones, pero su significado depende del análisis de los métodos, contextos y alcances de cada estudio.

La expansión de estas capacidades también ha hecho más visible un problema que atraviesa todo el proceso: la confiabilidad de las respuestas producidas por sistemas generativos. Walters y Wilder (2023) documentan referencias fabricadas y errores bibliográficos generados por ChatGPT, mientras Chelli et al. (2024) examinan problemas de alucinación y precisión de referencias en tareas relacionadas con revisiones sistemáticas. Una cita plausible, un DOI aparentemente válido o un resumen bien redactado no garantizan correspondencia con una fuente real. La verificación debe abarcar tanto los datos bibliográficos como la relación entre afirmación y evidencia.

Por esta razón, la incorporación de inteligencia artificial a la investigación científica plantea una exigencia transversal de trazabilidad. Debe ser posible reconstruir cómo se formuló una búsqueda, qué documentos fueron seleccionados, qué transformaciones experimentó la información y qué fuentes respaldan las afirmaciones elaboradas posteriormente. Penabad et al. (2024) destacan la necesidad de transparencia en el uso de inteligencia artificial dentro de los contextos científico-académicos. Esta exigencia convierte la documentación del proceso en

una condición para evaluar la reproducibilidad y confiabilidad de las actividades automatizadas.

Este libro examina estas transformaciones mediante cinco dimensiones vinculadas entre sí: búsqueda avanzada de literatura, construcción de mapas digitales del conocimiento, lectura automatizada de grandes colecciones, detección de tendencias y vacíos, y verificación inteligente de fuentes y referencias. El interés no reside en presentar la inteligencia artificial como sustituto del trabajo investigador, sino en determinar qué tareas puede apoyar, qué nuevas posibilidades analíticas introduce y qué controles necesita cada operación. La cuestión central es cómo ampliar la capacidad de explorar información sin debilitar la relación entre método, evidencia e interpretación.

La utilidad científica de estas herramientas depende, finalmente, de mantener una diferencia clara entre automatizar una operación y delegar una decisión. Un sistema puede localizar documentos, agrupar conceptos o señalar inconsistencias con una velocidad difícil de alcanzar manualmente; no por ello determina qué evidencia merece mayor peso ni qué interpretación resulta metodológicamente defendible. Trabajar con inteligencia artificial exige integrar capacidad computacional y juicio crítico dentro del mismo flujo de investigación. Esa relación define las nuevas formas de buscar y analizar información científica desarrolladas en esta obra.

Capítulo

1.

Búsqueda avanzada de literatura científica con inteligencia artificial

Wilmer Patricio Collaguazo Vega



1.1 Evolución de la búsqueda científica en entornos digitales

La búsqueda científica ha experimentado una transformación profunda a medida que los entornos digitales modificaron las formas de localizar, seleccionar y relacionar información académica. El tránsito desde catálogos y bases de datos centrados en coincidencias léxicas hacia sistemas capaces de interpretar relaciones semánticas amplió las posibilidades de recuperación documental. Gottardi et al. (2021) muestran que la búsqueda semántica responde precisamente a la necesidad de mejorar la localización de conocimiento científico mediante mecanismos que superen la correspondencia literal entre términos y documentos en contextos académicos complejos.

Este cambio no implica únicamente una sustitución tecnológica, sino una modificación del modo en que se formula y ejecuta la búsqueda. Los sistemas digitales tradicionales exigen una definición relativamente precisa de palabras clave, operadores y campos de consulta, mientras que las soluciones apoyadas en inteligencia artificial introducen procesos de interpretación contextual. Barradas (2023) sitúa la inteligencia artificial como un elemento transformador de la investigación científica, lo que permite comprender la búsqueda documental como una actividad cada vez más vinculada con procesos automatizados de clasificación, recomendación y análisis de información.

La expansión de la literatura científica también ha incrementado la dificultad de identificar documentos pertinentes dentro de grandes volúmenes de publicaciones. En este escenario, la búsqueda deja de consistir únicamente en recuperar registros y pasa a requerir mecanismos de priorización y filtrado. Mata et al. (2024), al revisar el uso de inteligencia artificial en la investigación científica, muestran la incorporación de estas tecnologías en distintas fases del trabajo investigativo. La recuperación de información se integra así en un flujo más amplio de apoyo computacional a la producción de conocimiento.

Los entornos digitales han favorecido, además, una relación más dinámica entre búsqueda, lectura y selección. Una consulta puede generar nuevos términos, autores relacionados, documentos vinculados y rutas de exploración que reorientan el proceso inicial. Cedeño et al. (2024) analizan la incorporación de inteligencia artificial en la investigación universitaria, contexto en el que estas herramientas adquieren funciones de apoyo para acceder, organizar y procesar información. La búsqueda científica comienza entonces a configurarse como un proceso iterativo, donde cada resultado puede modificar las decisiones posteriores del investigador.

La búsqueda semántica constituye uno de los cambios más relevantes dentro de esta evolución. A diferencia de los procedimientos basados exclusivamente en coincidencias textuales, busca recuperar documentos atendiendo a relaciones de significado y contexto. Gottardi et al. (2021) destacan su potencial para mejorar la encontrabilidad de la producción científica, especialmente cuando la terminología utilizada por quien consulta no coincide exactamente con la empleada en los documentos. Esta capacidad resulta útil en campos donde existen variaciones conceptuales, vocabularios especializados o formas distintas de denominar problemas semejantes.

La inteligencia artificial también ha favorecido nuevas formas de interacción con los sistemas de búsqueda. Las consultas pueden expresarse mediante lenguaje cercano a una pregunta de investigación, lo que reduce parcialmente la dependencia de estructuras rígidas de interrogación. Díaz (2024) examina el uso de inteligencia artificial en la investigación científica y advierte su incorporación creciente como herramienta de apoyo. Sin embargo, esta facilidad operativa no elimina la necesidad de formular con precisión el problema, porque una consulta ambigua puede conducir a resultados amplios, poco pertinentes o difíciles de evaluar críticamente.

La evolución digital también ha modificado la función del investigador frente a los resultados recuperados. Antes, gran parte del esfuerzo se concentraba en localizar documentos; ahora cobra mayor importancia evaluar la pertinencia de los materiales sugeridos por sistemas automatizados. Mujica (2021) vincula la alfabetización científica con las exigencias de la era digital, lo que refuerza la necesidad de comprender cómo se produce, selecciona y valida la información. El acceso más rápido a documentos no sustituye el juicio necesario para distinguir relevancia, calidad metodológica y utilidad para una investigación específica.

La automatización introduce, por tanto, una tensión entre eficiencia y control. Los sistemas pueden reducir el tiempo dedicado a explorar grandes conjuntos documentales, pero sus resultados dependen de criterios técnicos que no siempre son visibles para el usuario. Barradas (2023) permite situar esta cuestión dentro de la transformación amplia de la investigación científica por la inteligencia artificial. La eficiencia de una búsqueda no debería medirse solo por la cantidad o rapidez de los resultados, sino por su capacidad para aportar evidencia pertinente, rastreable y coherente con la pregunta formulada.

Otro cambio relevante corresponde a la integración entre búsqueda y gestión documental. Los resultados ya no se conciben únicamente como referencias aisladas, sino como conjuntos que

pueden organizarse, depurarse y relacionarse dentro de flujos digitales. Tapia y Recalde (2024), al revisar procesos de mejora de la gestión documental, aportan elementos para comprender la importancia de organizar información de manera sistemática y trazable. Esta integración favorece una continuidad entre localizar documentos, clasificarlos, eliminar duplicados y conservar aquellos que cumplen criterios definidos, aunque cada decisión exige supervisión metodológica.

La transformación digital también ha ampliado las posibilidades de acceso a literatura procedente de diferentes disciplinas y repositorios. Esta amplitud puede favorecer búsquedas más comprehensivas, pero también aumenta el riesgo de incorporar resultados heterogéneos sin suficiente discriminación. Cedeño et al. (2024) muestran que la inteligencia artificial se inserta en prácticas universitarias de investigación que requieren nuevas capacidades para gestionar información. Por ello, la evolución de la búsqueda científica no puede reducirse a la disponibilidad de herramientas; supone desarrollar criterios para delimitar fuentes, interpretar resultados y justificar decisiones de selección.

El uso de inteligencia artificial en la búsqueda científica plantea igualmente problemas vinculados con transparencia y supervisión. Castillo (2023) subraya la importancia de la intervención humana cuando ChatGPT se utiliza como apoyo en tareas científicas, principio que también resulta pertinente para la recuperación de literatura. Una recomendación automatizada puede ser útil como punto de partida, pero no debe aceptarse sin comprobación. La evolución hacia sistemas más autónomos exige, por tanto, mantener procedimientos de verificación capaces de identificar omisiones, errores, referencias inexistentes o relaciones documentales que carezcan de sustento suficiente.

La principal transformación no reside, entonces, en que la inteligencia artificial permita buscar más información, sino en que altera los criterios con los que se construye una búsqueda científicamente defendible. Un sistema puede ampliar términos, sugerir documentos o establecer relaciones entre contenidos, pero la calidad del proceso sigue dependiendo de decisiones metodológicas sobre pertinencia, cobertura y exclusión. La búsqueda avanzada exige conservar ese control porque una recuperación técnicamente sofisticada puede seguir siendo deficiente si no responde con precisión al problema investigado ni permite reconstruir cómo se seleccionó la evidencia.

1.2 Formulación inteligente de preguntas y estrategias de búsqueda

La formulación de una pregunta de búsqueda constituye una decisión metodológica, no una etapa meramente operativa. Cuando el problema científico está mal delimitado, incluso un sistema avanzado de recuperación puede devolver resultados extensos, pero poco útiles. Arias y Artigas (2022) vinculan la identificación de brechas con la capacidad de plantear problemas científicos relevantes, lo que permite entender que una buena búsqueda comienza antes de introducir términos en una plataforma. La calidad de la consulta depende de precisar qué se busca, bajo qué condiciones y con qué alcance conceptual.

Esa precisión exige traducir el problema de investigación a una estructura recuperable. En los entornos digitales, una pregunta demasiado amplia puede producir dispersión temática, mientras que una formulación excesivamente cerrada puede excluir documentos pertinentes. Gottardi et al. (2021) muestran que la búsqueda semántica amplía la posibilidad de recuperar información cuando existen diferencias entre el lenguaje utilizado por el investigador y el contenido de los documentos. Esto obliga a pensar la consulta como una construcción flexible que puede integrar términos equivalentes, relaciones conceptuales y distintas formas de expresar un mismo fenómeno.

La inteligencia artificial introduce un cambio adicional porque permite trabajar con consultas formuladas en lenguaje natural y con sistemas capaces de interpretar relaciones entre conceptos. Ajith et al. (2024), al proponer un referente para evaluar la recuperación de literatura científica, muestran que la búsqueda académica puede analizarse como un problema específico de recuperación, distinto de la búsqueda general en la web. La pertinencia de los resultados depende de que el sistema identifique documentos relevantes para una necesidad científica concreta, no solo textos que compartan palabras con la consulta inicial.

La formulación inteligente también requiere distinguir entre la pregunta científica y la estrategia de búsqueda. La primera delimita el conocimiento que se pretende obtener; la segunda traduce esa necesidad en decisiones sobre términos, filtros, relaciones y criterios de inclusión. Torre et al. (2023) analizan la automatización de revisiones sistemáticas mediante inteligencia artificial, mostrando que la búsqueda se integra con procesos posteriores de selección y análisis documental. Por ello, una estrategia eficaz debe diseñarse considerando desde el inicio cómo serán evaluados, filtrados y organizados los resultados recuperados.

El refinamiento progresivo ocupa un lugar central en este proceso. Una primera búsqueda puede revelar conceptos no previstos, denominaciones alternativas o líneas temáticas que obliguen a modificar la consulta inicial. Ofori et al. (2024) describen diferentes aplicaciones

del procesamiento del lenguaje natural, el aprendizaje automático y el aprendizaje profundo en la automatización de revisiones sistemáticas. Estas posibilidades convierten la búsqueda en un proceso iterativo: formular, recuperar, examinar, ajustar y volver a consultar hasta alcanzar un conjunto documental suficientemente pertinente para responder al problema planteado.

La intervención de sistemas inteligentes no elimina, sin embargo, la necesidad de justificar cada modificación de la estrategia. Una recomendación automática de términos puede resultar útil, pero debe evaluarse según su relación con el objeto de estudio. La ventaja técnica de ampliar una consulta pierde valor si incorpora conceptos periféricos o modifica silenciosamente el alcance original. Arias y Artigas (2022) permiten situar esta exigencia en el plano metodológico: identificar una brecha requiere distinguir qué conocimiento falta realmente y evitar que la amplitud de la búsqueda sustituya la delimitación conceptual del problema.

Tabla 1

Relación entre formulación de la pregunta y decisiones de búsqueda asistida por IA

Componente	Función metodológica	Riesgo si se formula de manera deficiente	Papel posible de la IA
Delimitación del problema	Define el fenómeno y el alcance de la búsqueda	Recuperación excesivamente amplia o irrelevante	Sugerir relaciones conceptuales y términos asociados
Conceptos principales	Traducen el problema a unidades recuperables	Omisión de literatura pertinente	Identificar variantes terminológicas y equivalencias semánticas
Criterios de inclusión	Determinan qué documentos responden a la necesidad científica	Incorporación de evidencia poco pertinente	Apoyar el filtrado y la priorización
Refinamiento iterativo	Ajusta la estrategia según los resultados obtenidos	Persistencia de sesgos o términos inadecuados	Proponer reformulaciones y nuevas rutas de consulta
Registro de decisiones	Permite reconstruir el proceso de búsqueda	Pérdida de trazabilidad y reproducibilidad	Facilitar documentación y organización del flujo de búsqueda

Nota. La tabla organiza las principales decisiones metodológicas que intervienen en la formulación y ajuste de búsquedas científicas asistidas por inteligencia artificial.

La relación mostrada en la tabla evidencia que la inteligencia artificial puede intervenir en distintas fases sin asumir la decisión final sobre la estrategia. Su aporte resulta más sólido cuando opera como soporte para ampliar, organizar o contrastar opciones, mientras el investigador conserva el control sobre el alcance científico de la consulta. Torre et al. (2023) y Ofori et al. (2024) muestran que la automatización adquiere mayor utilidad cuando forma parte

de un flujo estructurado, donde cada etapa puede ser revisada y vinculada con criterios previamente establecidos.

Otro aspecto decisivo es la capacidad de una estrategia para equilibrar sensibilidad y precisión. Recuperar demasiados documentos puede incrementar el esfuerzo de cribado, pero una búsqueda demasiado restrictiva puede dejar fuera estudios relevantes. Ajith et al. (2024) permiten comprender esta tensión desde la evaluación de sistemas de recuperación de literatura científica. La formulación inteligente no consiste en maximizar el número de resultados, sino en diseñar consultas que favorezcan la recuperación de evidencia pertinente y reduzcan, en la medida de lo posible, el ruido documental generado por relaciones conceptuales débiles.

La formulación de preguntas también debe conservar una conexión explícita con los criterios que posteriormente guiarán la selección de documentos. Si la consulta se modifica repetidamente sin registrar esos cambios, puede alterarse el universo documental sin que exista una justificación metodológica clara. Gottardi et al. (2021) muestran que la recuperación semántica amplía las posibilidades de encontrabilidad, pero esa amplitud exige mayor control sobre las decisiones del usuario. Una estrategia de búsqueda sólida debe permitir explicar por qué ciertos términos fueron incluidos, descartados o sustituidos durante el proceso.

El valor de una estrategia asistida por inteligencia artificial depende, finalmente, de su capacidad para conservar coherencia entre pregunta, consulta y evidencia recuperada. Una formulación técnicamente sofisticada puede resultar metodológicamente débil si el sistema orienta la búsqueda hacia temas próximos, pero distintos del problema original. El criterio central sigue siendo la correspondencia entre lo que se pretende conocer y los documentos seleccionados para responderlo. La inteligencia artificial puede ampliar las rutas de acceso a la literatura, pero la delimitación científica continúa siendo una responsabilidad metodológica que no conviene delegar.

1.3 Búsqueda semántica y recuperación contextual de información

La búsqueda semántica modifica el modo en que se recupera literatura científica porque desplaza parte del énfasis desde la coincidencia literal de palabras hacia la identificación de relaciones de significado. Gottardi et al. (2021) muestran que este enfoque busca mejorar la encontrabilidad de información científica mediante representaciones capaces de vincular consultas y documentos más allá de términos idénticos. Esta transformación resulta especialmente relevante cuando un mismo fenómeno aparece descrito con vocabularios

distintos, abreviaturas, conceptos relacionados o expresiones propias de comunidades disciplinares que no siempre comparten una terminología homogénea.

La recuperación contextual añade otra capa de complejidad, pues no basta con detectar que dos expresiones guardan similitud semántica. También interesa reconocer el contexto en el que un concepto aparece y la función que cumple dentro de un documento. Ajith et al. (2024) abordan la recuperación de literatura científica como una tarea que requiere identificar trabajos pertinentes para necesidades de información específicas. Bajo esta lógica, la calidad de una búsqueda depende menos de acumular coincidencias y más de recuperar documentos cuya relación con la consulta pueda justificarse por su contenido.

Los sistemas tradicionales basados en palabras clave continúan siendo útiles, pero presentan limitaciones cuando el investigador desconoce la terminología dominante de un campo o cuando una pregunta combina conceptos procedentes de áreas distintas. Gottardi et al. (2021) identifican precisamente la búsqueda semántica como una vía para favorecer el acceso a contenidos científicos cuya relevancia podría pasar inadvertida bajo mecanismos estrictamente léxicos. El cambio no elimina la función de las palabras clave; las incorpora dentro de un modelo de recuperación más amplio, donde significado, relaciones y contexto intervienen conjuntamente.

La integración de grafos de conocimiento amplía estas posibilidades al representar entidades y vínculos explícitos entre documentos, conceptos y otros elementos de la comunicación científica. Polo y Casique (2025) proponen una recuperación documental que combina grafos de conocimiento y redes neuronales, mostrando una alternativa para relacionar información estructurada con capacidades de aprendizaje automatizado. Este tipo de aproximación permite que una búsqueda no dependa únicamente de una consulta aislada, sino también de conexiones previamente identificadas entre conceptos, documentos y estructuras que forman parte de un dominio documental.

Los grafos académicos contribuyen, además, a contextualizar los resultados mediante relaciones que no siempre aparecen de manera evidente en el texto. Verma et al. (2023) examinan los grafos de conocimiento científico como estructuras destinadas a organizar la comunicación académica y representar entidades vinculadas entre sí. Desde el punto de vista de la recuperación, esta arquitectura permite explorar conexiones entre publicaciones, autores, conceptos y otros componentes del conocimiento científico. La utilidad reside en que la

pertinencia puede derivarse tanto del contenido documental como de las relaciones establecidas dentro de una red.

La recomendación académica comparte varios fundamentos con la búsqueda semántica, aunque responde a una lógica distinta. Mientras una consulta parte de una necesidad expresada por el usuario, los sistemas de recomendación pueden sugerir materiales a partir de similitudes, relaciones o patrones previamente identificados. Zhang et al. (2023) revisan los sistemas de recomendación académica y muestran la diversidad de enfoques empleados para proponer recursos científicos. Estas herramientas amplían la exploración documental, pero también requieren distinguir entre una recomendación potencialmente interesante y una evidencia realmente pertinente para el problema investigado.

La recuperación contextual resulta particularmente útil cuando la búsqueda se inicia con una pregunta compleja cuya respuesta no se encuentra expresada de forma literal en los títulos o resúmenes. Ajith et al. (2024) muestran la necesidad de evaluar sistemas específicamente orientados a literatura científica, donde la pertinencia depende de interpretar correctamente la necesidad planteada. Esta característica obliga a diferenciar entre similitud superficial y correspondencia científica. Dos documentos pueden compartir términos sin responder al mismo problema, mientras otros pueden emplear vocabularios diferentes y aportar evidencia directamente relacionada con la consulta.

La extracción de entidades y relaciones también desempeña un papel importante cuando se pretende convertir grandes colecciones documentales en estructuras que faciliten búsquedas más precisas. Jaradeh et al. (2023) analizan tuberías de extracción de información destinadas a construir grafos de conocimiento, proceso que implica identificar elementos significativos y establecer vínculos entre ellos. Aplicado a la literatura científica, este procedimiento favorece consultas basadas en relaciones explícitas y no únicamente en cadenas textuales. La recuperación adquiere así una dimensión estructural que permite navegar entre componentes conectados del conocimiento.

Sin embargo, una mayor capacidad para establecer conexiones no garantiza automáticamente mejores resultados. Las relaciones semánticas pueden introducir documentos próximos al tema, pero insuficientemente vinculados con la pregunta específica. Polo y Casique (2025) muestran que la recuperación documental apoyada en grafos y redes neuronales depende de una adecuada integración entre representación del conocimiento y procesamiento automatizado. Esto implica que el investigador debe examinar la calidad de las conexiones propuestas, evitando asumir

que toda proximidad semántica equivale a pertinencia metodológica, conceptual o empírica para una investigación determinada.

Figura 1

Evolución de la recuperación de información científica



Nota. La figura sintetiza la integración de enfoques basados en inteligencia artificial para ampliar la precisión y pertinencia en la recuperación de información científica.

La interpretación de los resultados exige, además, reconocer que los sistemas semánticos operan sobre representaciones construidas a partir de datos, modelos y relaciones previamente definidas. Brack et al. (2022), al analizar los requisitos de un grafo abierto de conocimiento de investigación, destacan la importancia de la calidad, la estructuración y las estrategias de construcción de estas representaciones. La recuperación contextual depende, por tanto, de la calidad del conocimiento que sustenta el sistema. Una estructura incompleta o sesgada puede limitar las conexiones disponibles y condicionar qué documentos aparecen como relevantes.

El valor metodológico de la búsqueda semántica reside en ampliar las rutas posibles hacia la evidencia sin convertir esa ampliación en una aceptación automática de los resultados. Cuando el sistema propone documentos por proximidad conceptual, relaciones de conocimiento o patrones de recomendación, todavía es necesario comprobar su correspondencia con la pregunta de investigación. La recuperación contextual funciona mejor como una forma de exploración guiada que como un mecanismo cerrado de selección. Su utilidad aumenta cuando permite descubrir conexiones no evidentes y, al mismo tiempo, conservar criterios explícitos para decidir cuáles resultan científicamente pertinentes.

1.4 Herramientas de inteligencia artificial para localizar literatura científica

Las herramientas de inteligencia artificial orientadas a la localización de literatura científica han ampliado las formas de acceder a documentos académicos, especialmente cuando el volumen de publicaciones dificulta una exploración manual exhaustiva. Su funcionamiento puede apoyarse en recuperación semántica, recomendación automatizada, procesamiento del lenguaje natural y modelos capaces de relacionar consultas con documentos relevantes. Zhang et al. (2023) muestran que los sistemas de recomendación académica emplean distintos enfoques para identificar recursos potencialmente útiles, lo que modifica la relación tradicional entre el investigador y las bases documentales.

Estas herramientas no cumplen toda la misma función. Algunas se concentran en recuperar documentos a partir de una consulta, mientras otras sugieren publicaciones relacionadas, priorizan resultados o facilitan la exploración de colecciones amplias. Ajith et al. (2024) examinan precisamente la recuperación de literatura científica como una tarea específica, donde la calidad del sistema depende de su capacidad para identificar trabajos relevantes ante necesidades de información complejas. Esta diferenciación resulta necesaria porque una plataforma que recomienda artículos no necesariamente ofrece las mismas garantías que un sistema diseñado para recuperar evidencia científica.

La incorporación de modelos de lenguaje ha permitido interactuar con ciertos sistemas mediante preguntas formuladas de manera más cercana al lenguaje natural. Esta posibilidad reduce algunas barreras técnicas asociadas con consultas rígidas, aunque no elimina la necesidad de formular adecuadamente el problema científico. Aiumtrakul et al. (2023), al comparar ChatGPT, BingChat y Bard en búsquedas de literatura relacionada con nefrología, muestran que estas herramientas pueden intervenir en la recuperación de información académica, pero también evidencian diferencias en precisión y confiabilidad de las referencias proporcionadas.

El potencial de estas tecnologías se observa con mayor claridad cuando se integran dentro de procesos de revisión de literatura. Torre et al. (2023) analizan la automatización de revisiones sistemáticas mediante inteligencia artificial y muestran que distintas etapas pueden recibir apoyo computacional, incluida la identificación inicial de documentos. Esto permite que la búsqueda se conecte con actividades posteriores de cribado y organización. La herramienta deja entonces de funcionar como un buscador aislado y pasa a formar parte de un flujo de trabajo que requiere decisiones metodológicas explícitas.

Una característica relevante de los sistemas inteligentes es su capacidad para reducir el esfuerzo dedicado a revisar resultados claramente alejados de la pregunta de investigación. Ofori et al. (2024) describen aplicaciones de procesamiento del lenguaje natural, aprendizaje automático y aprendizaje profundo en revisiones sistemáticas, donde la automatización puede apoyar distintas tareas documentales. Aplicada a la localización de literatura, esta capacidad permite priorizar registros potencialmente pertinentes. Sin embargo, la utilidad de esa priorización depende de los criterios con que se entrene o configure el sistema y del control ejercido durante su uso.

Los sistemas de recomendación académica introducen una lógica distinta a la consulta convencional porque pueden sugerir documentos sin que cada uno haya sido solicitado de manera explícita. Zhang et al. (2023) revisan múltiples aproximaciones utilizadas para recomendar artículos y otros recursos científicos, mostrando que estas herramientas pueden ampliar el espacio de exploración de una búsqueda. La ventaja consiste en descubrir conexiones que podrían pasar inadvertidas; el riesgo aparece cuando la recomendación se interpreta como sinónimo de pertinencia y se incorporan documentos simplemente porque el sistema los presenta como relacionados.

La comparación entre herramientas también exige valorar la procedencia y verificabilidad de los resultados. Aiumtrakul et al. (2023) muestran que las respuestas generadas por sistemas conversacionales pueden incluir referencias con distintos niveles de exactitud, lo que obliga a comprobar los datos bibliográficos antes de incorporarlos a una investigación. La rapidez con que una plataforma produce una lista de documentos puede generar una falsa impresión de seguridad. Localizar literatura científica requiere distinguir entre sugerencia, recuperación y validación, pues cada operación responde a criterios diferentes dentro del trabajo académico.

La eficiencia técnica adquiere mayor utilidad cuando se articula con una estrategia de búsqueda definida previamente. Ajith et al. (2024) evidencian que la recuperación científica puede evaluarse según la capacidad de localizar documentos relevantes frente a determinadas necesidades de información. Esto implica que una herramienta no debería seleccionarse únicamente por la facilidad de uso o por la cantidad de resultados obtenidos. Su pertinencia depende de cómo responde al tipo de consulta planteada, del campo disciplinar involucrado y del nivel de precisión requerido para construir un conjunto documental adecuado.

En el ámbito universitario, estas herramientas también están modificando prácticas vinculadas con la búsqueda y organización de información científica. Cedeño et al. (2024) analizan el uso

de inteligencia artificial en procesos de investigación universitaria y muestran su incorporación como recurso de apoyo en actividades académicas. Esta presencia puede facilitar el acceso inicial a literatura, pero también exige competencias para interpretar los resultados obtenidos. El investigador debe reconocer cuándo una plataforma está recuperando fuentes reales, cuándo está generando respuestas sintéticas y cuándo ambas funciones aparecen combinadas dentro de una misma interfaz.

La automatización de revisiones permite observar otra diferencia importante entre herramientas generales y herramientas especializadas. Torre et al. (2023) estudian sistemas diseñados para apoyar procedimientos estructurados de revisión, mientras Ofori et al. (2024) muestran la diversidad de técnicas computacionales aplicables a esas tareas. Una herramienta especializada puede ofrecer funciones orientadas al cribado o clasificación documental que no están presentes en asistentes conversacionales de propósito general. Elegir entre ambas opciones exige considerar la finalidad concreta de la búsqueda y el nivel de control metodológico requerido.

La localización asistida por inteligencia artificial tampoco debe evaluarse únicamente por su capacidad para ahorrar tiempo. Un sistema puede recuperar rápidamente numerosos documentos y, al mismo tiempo, dejar fuera trabajos relevantes o incluir registros cuya relación con el problema sea débil. Aiumtrakul et al. (2023) muestran que la confiabilidad de las referencias obtenidas mediante sistemas generativos requiere comprobación independiente. Este punto obliga a conservar una separación clara entre velocidad de recuperación y calidad de la evidencia, especialmente cuando los resultados serán utilizados para sustentar argumentos o decisiones investigativas.

La elección de una herramienta adquiere valor científico cuando responde a una necesidad concreta y puede integrarse dentro de un procedimiento verificable. No resulta metodológicamente suficiente utilizar una plataforma porque genere respuestas extensas, sugiera artículos o simplifique consultas. Lo decisivo es establecer qué función cumple dentro del proceso, qué tipo de resultados produce y cómo serán comprobados antes de incorporarlos al corpus de estudio. La inteligencia artificial puede ampliar las posibilidades de localización, pero su utilidad depende de que la herramienta se subordine a criterios de pertinencia, trazabilidad y control documental.

1.5 Refinamiento, filtrado y priorización de resultados

La recuperación inicial de literatura rara vez produce un conjunto documental listo para ser utilizado de forma inmediata. Una búsqueda amplia puede incluir registros duplicados, estudios periféricos, documentos con escasa relación con la pregunta y publicaciones que solo comparten parte del vocabulario empleado. El refinamiento cumple la función de transformar ese universo inicial en un corpus más pertinente y manejable. Dennstädt et al. (2024) muestran que los modelos de lenguaje pueden apoyar el cribado de títulos y resúmenes, lo que introduce nuevas posibilidades para reducir la carga asociada con la revisión preliminar de resultados.

Filtrar no equivale únicamente a eliminar documentos. También supone aplicar criterios explícitos que permitan justificar por qué determinados registros permanecen dentro del conjunto analizado y otros son excluidos. Torre et al. (2023) sitúan la automatización dentro de distintas fases de las revisiones sistemáticas, mostrando que la inteligencia artificial puede intervenir en tareas de selección sin sustituir la definición metodológica previa. El valor del filtrado depende, por tanto, de que los criterios de inclusión y exclusión respondan al problema científico y puedan mantenerse de manera consistente durante el proceso.

La priorización introduce una operación diferente. Mientras el filtrado decide qué documentos cumplen condiciones mínimas, priorizar implica establecer un orden de revisión entre aquellos que permanecen. Ajith et al. (2024) permiten comprender esta tarea desde la lógica de la recuperación científica, donde los sistemas deben distinguir documentos más relevantes frente a una necesidad de información concreta. La posición de un resultado adquiere importancia cuando el volumen recuperado es elevado, porque los primeros registros suelen recibir mayor atención y pueden influir tempranamente en la comprensión que el investigador construye sobre el campo.

Los procedimientos automatizados pueden utilizar características textuales y relaciones aprendidas para apoyar esa clasificación. Ofori et al. (2024) examinan técnicas de procesamiento del lenguaje natural, aprendizaje automático y aprendizaje profundo aplicadas a la automatización de revisiones sistemáticas. Su utilidad reside en identificar patrones que permitan reducir el trabajo repetitivo asociado con grandes volúmenes documentales. Sin embargo, una clasificación automatizada sigue dependiendo de las señales que utiliza el modelo, por lo que un resultado colocado en una posición inferior no debe interpretarse automáticamente como irrelevante para la investigación.

La organización documental también influye en la calidad del refinamiento. Tapia y Recalde (2024) destacan, desde la gestión documental, la necesidad de estructurar procesos que faciliten

el tratamiento sistemático de la información. En una búsqueda científica, esta lógica se traduce en conservar registros ordenados, reconocer duplicados y mantener información suficiente para reconstruir las decisiones tomadas. El filtrado asistido por IA resulta más útil cuando se integra con procedimientos de gestión que eviten perder documentos, aplicar criterios de manera desigual o modificar el corpus sin dejar evidencia del cambio realizado.

Tabla 2

Relación entre operaciones de depuración y decisiones de selección documental

Operación	Pregunta que orienta la decisión	Aporte posible de la IA	Riesgo metodológico
Depuración	¿El registro es único y utilizable?	Detectar duplicados y organizar registros	Eliminar versiones relevantes por similitud aparente
Filtrado	¿Cumple los criterios establecidos?	Clasificar títulos y resúmenes según pertinencia	Reproducir criterios insuficientemente definidos
Priorización	¿Qué documentos deben revisarse primero?	Ordenar registros según probabilidad de relevancia	Confundir posición algorítmica con calidad científica
Revisión de exclusiones	¿Los descartes fueron adecuados?	Señalar casos ambiguos o fronterizos	Aceptar exclusiones automáticas sin comprobación
Registro del proceso	¿Puede reconstruirse la selección?	Facilitar organización y trazabilidad documental	Depender de decisiones no documentadas del sistema

Nota. La tabla diferencia operaciones que suelen confluir durante la reducción de resultados y muestra el tipo de control necesario cuando intervienen sistemas automatizados.

La distinción entre estas operaciones evita tratar el proceso como una única decisión automática. Un documento puede superar el filtrado inicial y, aun así, recibir una prioridad baja; del mismo modo, un registro descartado automáticamente puede requerir revisión cuando presenta características ambiguas. Dennstädt et al. (2024) muestran el potencial de los modelos de lenguaje para apoyar el cribado, pero la aplicación práctica exige conservar mecanismos para examinar casos dudosos. La eficiencia se vuelve metodológicamente útil cuando reduce trabajo repetitivo sin cerrar prematuramente decisiones que todavía requieren juicio investigador.

La calidad del filtrado también depende de la estabilidad de los criterios empleados. Si las reglas cambian después de observar ciertos resultados, la selección puede terminar respondiendo a expectativas surgidas durante la revisión y no a condiciones previamente justificadas. Torre et al. (2023) permiten situar la automatización dentro de un flujo estructurado, donde las distintas fases deben conservar coherencia. Registrar ajustes, revisar su

efecto y aplicarlos de forma uniforme ayuda a evitar que la flexibilidad propia de una búsqueda asistida por IA se convierta en una fuente adicional de arbitrariedad documental.

Priorizar resultados exige además reconocer que relevancia temática y calidad científica no son equivalentes. Un artículo puede responder de manera directa a la consulta y, sin embargo, presentar limitaciones que reduzcan su utilidad como evidencia. Ajith et al. (2024) se concentran en el problema de recuperar literatura pertinente, lo que permite separar esa tarea de la evaluación posterior del contenido. Los sistemas de ranking ayudan a decidir qué revisar primero, pero no deberían determinar por sí solos qué publicaciones merecen mayor peso dentro de una argumentación científica.

El refinamiento resulta más sólido cuando reduce progresivamente el conjunto documental sin ocultar las decisiones que hicieron posible esa reducción. La automatización puede acelerar la identificación de duplicados, apoyar el cribado y ordenar registros, pero una selección defendible necesita conservar la posibilidad de revisar exclusiones y reconstruir criterios. El objetivo no consiste en obtener el menor número posible de documentos, sino en formar un corpus suficientemente concentrado para ser analizado y suficientemente amplio para no perder evidencia relevante por decisiones automáticas difíciles de justificar.

1.6 Evaluación de pertinencia y calidad de los documentos recuperados

La recuperación de un documento no garantiza que su contenido sea adecuado para responder una pregunta científica. La pertinencia se relaciona con el grado en que una publicación aborda el problema, las variables, la población o el contexto definidos en la búsqueda, mientras que la calidad exige examinar características adicionales del trabajo recuperado. Ajith et al. (2024) muestran que los sistemas de recuperación científica deben evaluarse por su capacidad para identificar literatura relevante, lo que obliga a distinguir entre coincidencia temática, utilidad metodológica y valor efectivo para una investigación determinada.

La pertinencia debe valorarse en relación con la pregunta que originó la búsqueda, no con la mera presencia de términos coincidentes. Un artículo puede contener las palabras empleadas en una consulta y, sin embargo, estudiar un fenómeno distinto o responder a otro propósito. Aiumtrakul et al. (2023), al analizar herramientas generativas empleadas para localizar literatura, evidencian que los resultados obtenidos requieren comprobación antes de incorporarse al trabajo académico. Esta evaluación evita que la amplitud de una respuesta automatizada se confunda con una selección documental rigurosa.

La calidad plantea un problema diferente porque un documento temáticamente pertinente puede presentar debilidades que reduzcan su utilidad como evidencia. De la Cruz (2017) relaciona el rigor científico con exigencias que afectan la consistencia de una investigación, lo que permite trasladar esa preocupación al examen de los documentos recuperados. La valoración debe considerar si el trabajo ofrece información suficiente para comprender cómo se construyeron sus resultados y si existe correspondencia entre objetivos, procedimientos y argumentos. La pertinencia temática, por sí sola, no resuelve estas cuestiones.

Los sistemas basados en inteligencia artificial pueden apoyar la identificación inicial de documentos potencialmente útiles, aunque esa asistencia no equivale a una evaluación definitiva de calidad. Chelli et al. (2024) comparan la precisión de referencias proporcionadas por herramientas generativas en tareas vinculadas con revisiones sistemáticas y muestran que pueden presentarse problemas de exactitud. La comprobación debe incluir la existencia real de la fuente y la correspondencia de sus datos bibliográficos antes de analizar su contenido. Una referencia plausible no debe asumirse automáticamente como una publicación válida.

La evaluación también requiere distinguir entre calidad de la fuente y calidad de la respuesta generada a partir de ella. Mugaanyi et al. (2024) estudian la fiabilidad de modelos de lenguaje en la producción de citas y referencias académicas y muestran la necesidad de examinar su precisión bibliográfica. Cuando una herramienta recomienda una publicación, el investigador debe regresar al documento original y no limitarse a la descripción proporcionada por el sistema. Esta práctica preserva la relación directa entre la interpretación realizada y la evidencia efectivamente contenida en la fuente consultada.

Tabla 3
Dimensiones para evaluar documentos recuperados mediante búsquedas asistidas por IA

Dimensión	Pregunta de evaluación	Evidencia que debe revisarse	Riesgo de una valoración insuficiente
Pertinencia temática	¿El documento responde al problema investigado?	Objetivo, tema, variables y contexto	Incorporar estudios próximos, pero no relevantes
Correspondencia metodológica	¿El diseño aporta evidencia adecuada para la pregunta?	Método, muestra, procedimientos y análisis	Utilizar resultados que no responden al tipo de evidencia requerida
Trazabilidad bibliográfica	¿La referencia puede comprobarse?	Autoría, título, revista, año y DOI	Citar documentos inexistentes o datos incorrectos

Dimensión	Pregunta de evaluación	Evidencia que debe revisarse	Riesgo de una valoración insuficiente
Consistencia interna	¿Existe relación entre objetivos, método y resultados?	Estructura argumentativa y reporte del estudio	Dar peso a conclusiones insuficientemente sustentadas
Utilidad para el corpus	¿Aporta información necesaria para el análisis?	Contribución específica frente a otros documentos	Acumular publicaciones redundantes

Nota. La tabla organiza criterios diferenciados para evitar que la relevancia temática se utilice como único indicador de selección documental.

La separación entre dimensiones permite evitar evaluaciones demasiado rápidas. Un documento puede cumplir con el tema de búsqueda y, al mismo tiempo, aportar escasa información para el objetivo específico del estudio. Castillo (2023) subraya la necesidad de supervisión humana en el uso de ChatGPT como apoyo científico, criterio que resulta aplicable al proceso de selección documental. La herramienta puede señalar coincidencias o sugerir fuentes, pero corresponde al investigador examinar si cada publicación añade evidencia sustantiva y si su incorporación al corpus está metodológicamente justificada.

Otro aspecto necesario es detectar la falsa apariencia de autoridad que puede producir una referencia formalmente convincente. Aiumtrakul et al. (2023) muestran que las herramientas generativas pueden ofrecer referencias cuya exactitud debe ser verificada, mientras Mugaanyi et al. (2024) examinan problemas de fiabilidad bibliográfica en distintos ámbitos disciplinares. Por ello, el formato correcto de una cita no constituye prueba de existencia ni de calidad. La comprobación debe dirigirse al documento original y a sus metadatos, separando presentación bibliográfica de validez efectiva.

La evaluación de calidad tampoco debería transformarse en una exclusión automática basada únicamente en indicadores externos. De la Cruz (2017) sitúa el rigor en la coherencia y solidez del proceso investigativo, lo que exige revisar el contenido del estudio y no reducir la valoración a señales superficiales. En la práctica, esto supone examinar cómo se planteó el problema, qué procedimientos se utilizaron y hasta dónde llegan las conclusiones. Un documento relevante puede ser útil incluso para identificar limitaciones, siempre que esas limitaciones sean reconocidas durante su incorporación al análisis.

Cuando la inteligencia artificial interviene en la selección, resulta especialmente necesario revisar los casos fronterizos. Un sistema puede clasificar como poco pertinente un estudio cuyo aporte aparece en una sección específica o cuya terminología difiere de la consulta inicial. Ajith et al. (2024) muestran que la recuperación científica plantea dificultades particulares

precisamente porque la relevancia no siempre se expresa mediante coincidencias evidentes. La evaluación humana conserva valor en estos casos, pues permite interpretar relaciones conceptuales, metodológicas o empíricas que un mecanismo automatizado puede ponderar de manera insuficiente.

La decisión de conservar un documento debe poder explicarse mediante criterios identificables y no por la posición que ocupó en una lista generada automáticamente. Un corpus sólido se construye cuando cada fuente cumple una función reconocible respecto del problema investigado y cuando sus datos pueden verificarse directamente. Este criterio obliga a separar comodidad tecnológica de juicio académico. La inteligencia artificial puede facilitar la exploración y reducir parte del trabajo inicial, pero la incorporación definitiva de una publicación requiere una lectura que determine qué aporta, qué limita y por qué merece formar parte de la evidencia analizada.

1.7 Diseño de estrategias reproducibles de búsqueda asistida por IA

Una estrategia de búsqueda reproducible debe permitir que otro investigador comprenda cómo se construyó el conjunto documental y, en condiciones semejantes, pueda repetir el procedimiento. Esta exigencia se vuelve más compleja cuando intervienen herramientas de inteligencia artificial, porque algunas operaciones pueden depender de modelos, configuraciones o recomendaciones dinámicas. Torre et al. (2023) muestran que la automatización de revisiones sistemáticas puede abarcar diversas etapas del proceso, por lo que documentar las decisiones de búsqueda deja de ser una tarea secundaria y pasa a formar parte de la calidad metodológica del estudio.

La reproducibilidad comienza con una formulación suficientemente explícita de la consulta. Deben registrarse los términos utilizados, sus combinaciones, filtros aplicados, fuentes consultadas y criterios empleados para modificar la búsqueda inicial. Ofori et al. (2024) analizan la incorporación de procesamiento del lenguaje natural, aprendizaje automático y aprendizaje profundo en revisiones sistemáticas, evidenciando que la automatización puede intervenir en distintas decisiones documentales. Cuando estas decisiones no quedan registradas, resulta difícil establecer si los resultados obtenidos dependen del diseño metodológico o de comportamientos específicos del sistema utilizado.

El uso de inteligencia artificial añade la necesidad de documentar cómo intervino la herramienta. No es suficiente indicar que se utilizó un asistente o un sistema automatizado; interesa precisar qué función cumplió dentro del proceso. Penabad et al. (2024) plantean

criterios para el uso y reporte de inteligencia artificial en contextos científico-académicos, reforzando la importancia de transparentar su participación. En una búsqueda, esto implica distinguir si la IA propuso términos, amplió consultas, priorizó documentos, eliminó registros o intervino en la selección final del corpus.

Otro componente esencial es conservar la secuencia de modificaciones realizadas durante la búsqueda. Las consultas suelen cambiar a medida que aparecen nuevos conceptos, variantes terminológicas o resultados insuficientes, pero esos ajustes deben quedar documentados. Tapia y Recalde (2024) destacan la importancia de una gestión documental organizada, principio que puede trasladarse al registro de las decisiones de búsqueda. Mantener versiones de consultas y criterios permite reconstruir cómo se pasó de una estrategia inicial a otra más refinada, evitando que las modificaciones queden reducidas a decisiones informales difíciles de justificar.

La reproducibilidad también depende de la estabilidad de los datos y herramientas utilizadas. Kapoor y Narayanan (2023) advierten sobre problemas de reproducibilidad en investigaciones basadas en aprendizaje automático, especialmente cuando existen filtraciones de información o condiciones experimentales insuficientemente controladas. Aunque su análisis se concentra en ciencia basada en aprendizaje automático, el principio resulta pertinente para búsquedas asistidas por IA: si el sistema cambia, si su índice se actualiza o si sus criterios internos no son transparentes, una misma consulta puede producir resultados distintos en momentos diferentes.

Por esa razón, registrar la fecha y el entorno de ejecución adquiere relevancia metodológica. Una búsqueda realizada con una plataforma dinámica no puede tratarse como si utilizara una base documental completamente estática. Semmelrock et al. (2025) examinan barreras y factores asociados con la reproducibilidad en investigaciones basadas en aprendizaje automático, mostrando que la documentación de procesos y condiciones técnicas resulta indispensable. En búsquedas científicas asistidas por IA, esta lógica exige conservar información suficiente sobre el sistema utilizado y las condiciones bajo las cuales fueron generados los resultados.

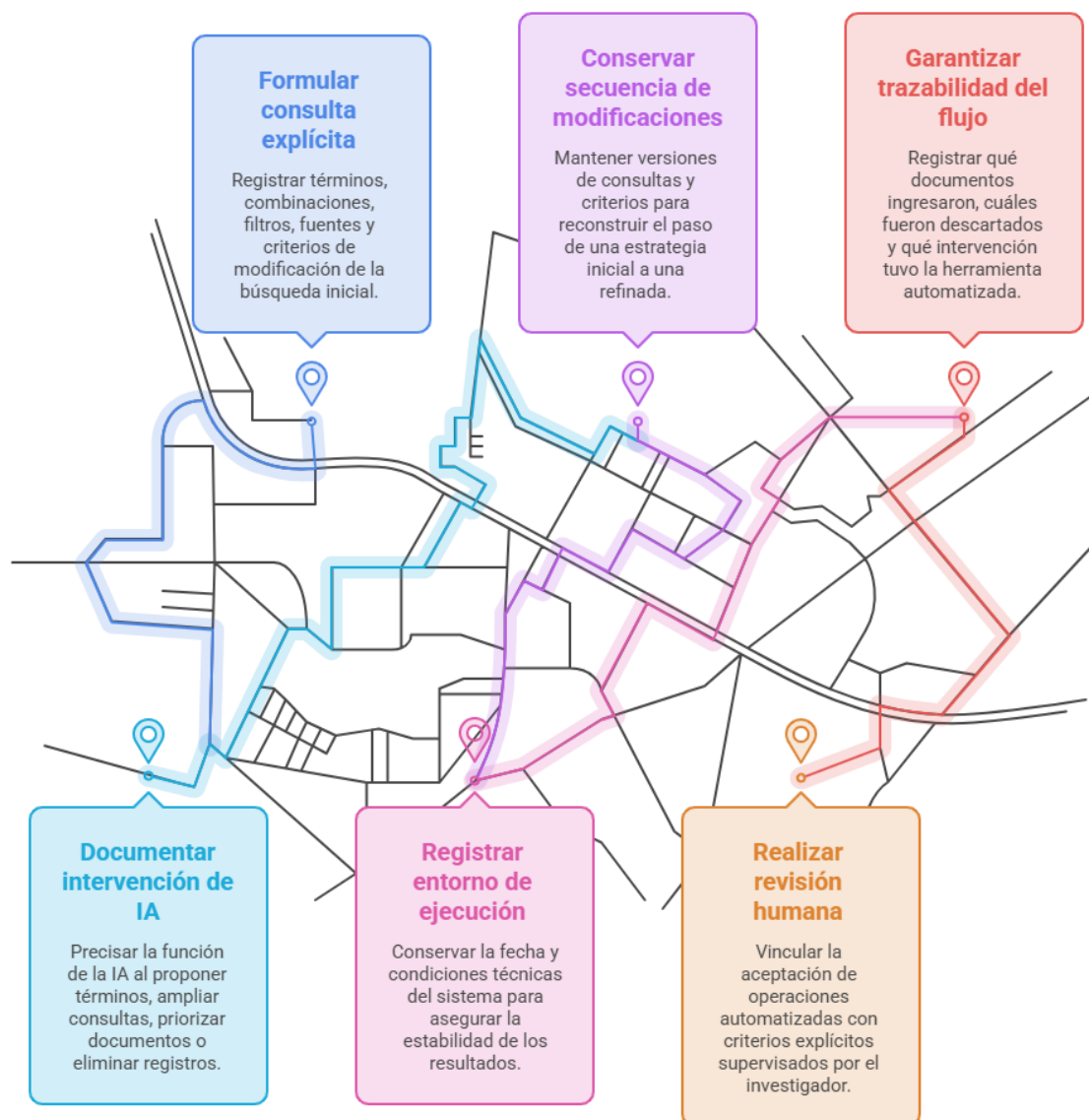
La trazabilidad del flujo de trabajo permite, además, identificar en qué momento se produjo una exclusión o una priorización determinada. Loza et al. (2026) abordan la reproducibilidad de flujos bioinformáticos y resaltan la importancia de estructurar procedimientos que puedan ser reconstruidos. Traslado a la recuperación documental, esto significa registrar qué documentos ingresaron inicialmente, cuáles fueron descartados, con qué criterio y qué

intervención tuvo la herramienta automatizada. Sin esta secuencia, la selección final puede parecer metodológicamente ordenada, aunque el proceso real haya sido difícil de reconstruir.

Una estrategia reproducible también debe prever la revisión humana de decisiones automatizadas. La IA puede facilitar el cribado o sugerir modificaciones, pero la aceptación de esas operaciones debe quedar vinculada con criterios explícitos. Torre et al. (2023) muestran que la automatización puede apoyar distintas fases de una revisión, aunque ello no elimina la necesidad de estructurar el procedimiento. La supervisión permite detectar casos en los que el sistema interpreta de manera insuficiente una consulta o descarta documentos que cumplen los criterios científicos establecidos por el investigador.

Figura 2

Diseño de estrategias reproducibles de búsqueda asistida por IA



Nota. La figura organiza los componentes necesarios para documentar, verificar y reproducir una estrategia de búsqueda científica asistida por inteligencia artificial.

La posibilidad de repetir una búsqueda no implica obtener exactamente el mismo conjunto de resultados en todos los casos. En sistemas dinámicos, pueden variar los índices, los modelos o la disponibilidad documental. Semmelrock et al. (2025) permiten comprender que la reproducibilidad también depende de describir adecuadamente el contexto de ejecución y no solo de conservar una secuencia de comandos. Por ello, una estrategia bien documentada debe facilitar la comparación entre resultados obtenidos en distintos momentos y permitir identificar si las diferencias provienen del procedimiento o de cambios externos en la herramienta.

La reproducibilidad también cumple una función crítica frente al sesgo de selección. Cuando las decisiones sobre inclusión, exclusión o priorización no están registradas, resulta más difícil detectar si el corpus fue construido de manera consistente. Penabad et al. (2024) refuerzan la necesidad de transparencia en el uso de inteligencia artificial dentro de la actividad científica. Aplicado a la búsqueda, esto implica hacer visible qué parte del proceso fue automatizada y qué decisiones permanecieron bajo control humano, evitando presentar como neutral una selección condicionada por sistemas cuya lógica puede no ser completamente observable.

Una estrategia reproducible de búsqueda asistida por IA debe conservar suficiente evidencia para reconstruir el camino seguido desde la pregunta inicial hasta el corpus definitivo. Esa exigencia no busca convertir cada búsqueda en un procedimiento rígido, sino impedir que la flexibilidad tecnológica vuelva opacas las decisiones metodológicas. Registrar consultas, cambios, criterios, fechas, herramientas y revisiones humanas permite evaluar la solidez del proceso incluso cuando los resultados varían. La reproducibilidad funciona, así, como un mecanismo de control frente a una automatización cuya utilidad depende de poder ser examinada y discutida.

Capítulo

2.

Construcción de mapas digitales del conocimiento científico

Cindy Estefania Alvia Toala



2.1 Organización del conocimiento científico mediante representaciones digitales

La organización del conocimiento científico mediante representaciones digitales responde a la necesidad de estructurar información que, por su volumen y complejidad, resulta difícil de comprender como una sucesión aislada de documentos. Las publicaciones pueden vincularse por conceptos, autores, métodos, instituciones o relaciones de citación, generando estructuras que facilitan su interpretación conjunta. Acevedo et al. (2017) destacan que los modelos científicos cumplen funciones de representación y comprensión, principio que puede trasladarse al entorno digital cuando se construyen esquemas capaces de mostrar relaciones relevantes dentro de un campo de conocimiento.

Las representaciones digitales no reproducen de forma neutral la totalidad de la producción científica. Toda estructura selecciona determinados elementos y deja otros en segundo plano según el criterio utilizado para organizar la información. Ávila (2024) muestra que los grafos de conocimiento pueden emplearse para visualizar datos abiertos de investigación mediante relaciones estructuradas entre entidades. Esta forma de representación permite superar una lectura basada únicamente en registros independientes, porque introduce conexiones explícitas que ayudan a observar cómo distintos componentes de la producción científica se relacionan dentro de un mismo sistema documental.

Un grafo de conocimiento puede representar entidades científicas como nodos y establecer entre ellas relaciones que aportan sentido a la estructura. Brack et al. (2022) analizan los requisitos para construir un grafo abierto de conocimiento de investigación y muestran que su utilidad depende tanto de los casos de uso como de la calidad de los datos y las estrategias de construcción. La representación digital adquiere valor cuando las conexiones expresadas permiten responder preguntas concretas y no cuando se limita a acumular enlaces sin una función analítica claramente definida.

La estructuración del conocimiento también depende de la capacidad para reconocer entidades y relaciones dentro de documentos originalmente no estructurados. Jaradeh et al. (2023) estudian procesos de extracción de información destinados a construir grafos de conocimiento, lo que permite comprender que la representación digital requiere transformar contenido textual en elementos organizables. Esta operación implica identificar conceptos, atributos y vínculos que puedan ser procesados posteriormente. El paso del texto al grafo modifica la forma de explorar la literatura, pues convierte fragmentos documentales dispersos en relaciones susceptibles de comparación.

Verma et al. (2023) sitúan los grafos de conocimiento académico dentro de la organización de la comunicación científica, donde publicaciones, autores y conceptos pueden representarse como componentes interconectados. Esta estructura favorece una lectura relacional del campo científico, ya que permite observar no solo qué documentos existen, sino también cómo se vinculan. Sin embargo, la utilidad del grafo depende de la coherencia con que se definan las entidades y relaciones. Una representación excesivamente simplificada puede ocultar diferencias conceptuales importantes entre documentos aparentemente próximos.

Tabla 4

Dimensiones analíticas de la organización digital del conocimiento científico

Dimensión	Elemento representado	Función analítica	Riesgo de representación
Entidades	Autores, documentos, conceptos, instituciones	Identificar componentes relevantes del campo	Omitir entidades necesarias para interpretar el corpus
Relaciones	Citación, autoría, afinidad conceptual, pertenencia	Mostrar conexiones entre componentes	Representar vínculos débiles como relaciones sustantivas
Jerarquías	Categorías, temas y niveles de organización	Ordenar el conocimiento por niveles	Imponer clasificaciones rígidas a fenómenos complejos
Atributos	Fecha, disciplina, tipo documental, procedencia	Diferenciar características de los nodos	Sobrevalorar metadatos frente al contenido
Estructura global	Red completa de relaciones	Reconocer patrones y agrupamientos	Confundir densidad visual con relevancia científica

Nota. La tabla organiza los principales componentes que intervienen en una representación digital y los riesgos asociados con su interpretación.

La utilidad de estas dimensiones radica en que obligan a separar la estructura técnica de la interpretación científica. Un grafo puede mostrar múltiples conexiones, pero la presencia de una relación no determina por sí misma su importancia. Ávila (2024) evidencia que la visualización de datos de investigación puede apoyarse en modelos formales de representación, mientras Brack et al. (2022) subrayan la necesidad de considerar criterios de calidad. El análisis debe examinar qué representa cada vínculo, con qué datos fue construido y qué conclusiones pueden extraerse razonablemente de su configuración.

Las representaciones digitales permiten integrar fuentes heterogéneas dentro de una estructura común, siempre que existan criterios compatibles para describirlas. Verma et al. (2023) muestran que los grafos académicos pueden organizar diversos componentes de la comunicación científica, mientras Jaradeh et al. (2023) explican cómo la extracción de información contribuye a convertir contenido documental en estructuras relacionables. Esta

integración facilita observar conexiones que serían difíciles de identificar mediante lectura lineal, aunque también exige controlar ambigüedades terminológicas y diferencias entre formas de describir un mismo objeto científico.

La representación mediante grafos ofrece una ventaja adicional: permite cambiar la escala de observación. Un investigador puede examinar relaciones específicas entre dos entidades o desplazarse hacia configuraciones más amplias del campo. Brack et al. (2022) muestran que los usos de un grafo de conocimiento dependen de las preguntas que se desean responder. Por ello, la estructura no debería diseñarse como una visualización genérica, sino como un sistema capaz de facilitar distintos niveles de análisis, desde relaciones locales hasta patrones que involucren conjuntos extensos de producción científica.

La organización digital también puede influir en la forma en que se percibe un campo científico. Una representación visual tiende a destacar agrupamientos, nodos centrales o relaciones densas, lo que puede inducir interpretaciones rápidas sobre relevancia o influencia. Acevedo et al. (2017) permiten recordar que toda representación científica implica decisiones sobre qué aspectos del fenómeno se consideran significativos. Aplicado a los mapas digitales, esto exige diferenciar entre propiedades del conocimiento representado y efectos producidos por la forma específica en que los datos fueron estructurados y visualizados.

El valor de una representación digital depende, finalmente, de que permita formular preguntas mejores sobre el conocimiento organizado. Un mapa útil no se limita a mostrar conexiones atractivas, sino que ayuda a identificar relaciones que merecen ser examinadas con mayor profundidad. La organización mediante entidades, atributos y vínculos puede facilitar esa exploración, pero no reemplaza la lectura del contenido ni la evaluación de la evidencia. La representación debe funcionar como una herramienta analítica que oriente la investigación sin convertir la estructura visual en una explicación suficiente del campo científico.

2.2 Redes de autores, documentos y conceptos científicos

Las redes científicas permiten representar la producción académica como una estructura de relaciones y no únicamente como una colección de publicaciones independientes. En este tipo de organización, autores, documentos y conceptos pueden funcionar como nodos conectados por vínculos de distinta naturaleza. Ávila (2024) muestra que los grafos de conocimiento aplicados a datos abiertos de investigación facilitan la visualización de relaciones entre entidades científicas. Esta lógica resulta útil para examinar cómo se articula un campo, qué

elementos concentran conexiones y qué zonas aparecen menos integradas dentro de la estructura general.

Las redes de autores suelen construirse a partir de vínculos de colaboración, participación conjunta en publicaciones o pertenencia a determinadas comunidades académicas. Haunschild y Bornmann (2024), al trabajar con OpenAlex para generar mapas bibliométricos globales, muestran el potencial de los datos abiertos para analizar estructuras científicas en distintos niveles. Estas representaciones permiten observar patrones de concentración y conexión que no son evidentes en una lista bibliográfica convencional. Sin embargo, la existencia de numerosos vínculos no debe interpretarse automáticamente como mayor calidad o influencia científica.

Las redes documentales se organizan a partir de relaciones entre publicaciones, como citas, similitud temática o conexiones derivadas de metadatos compartidos. Verma et al. (2023) explican que los grafos de conocimiento académico integran distintos componentes de la comunicación científica dentro de estructuras relacionables. Esta perspectiva permite seguir trayectorias entre documentos y reconocer agrupamientos que sugieren proximidad conceptual o bibliográfica. El análisis cobra mayor valor cuando las conexiones se interpretan junto con el contenido de los trabajos y no únicamente a partir de su posición dentro de la red.

Los conceptos científicos conforman un tercer nivel de organización que puede complementar las redes de autores y documentos. Una misma publicación puede vincular varios conceptos, mientras un concepto puede aparecer distribuido entre numerosos trabajos y comunidades de investigación. Brack et al. (2022) muestran que la construcción de grafos abiertos de conocimiento exige definir con precisión las entidades y relaciones que se desean representar. Esta exigencia resulta especialmente importante en el caso de conceptos, porque términos similares pueden tener significados diferentes según el campo disciplinar o el contexto de uso.

La relación entre autores, documentos y conceptos permite construir representaciones más ricas que aquellas centradas en una sola dimensión. Un autor puede conectarse con publicaciones específicas y, a través de ellas, con temas, métodos o problemas de investigación. Ávila (2024) muestra que los grafos pueden integrar distintos tipos de datos y facilitar una lectura relacional de la producción científica. Esta articulación ayuda a identificar no solo quién publica y qué publica, sino también qué ideas circulan entre comunidades y qué conexiones temáticas se forman alrededor de determinados conjuntos documentales.

La utilidad analítica de estas redes depende en gran medida de la calidad de los datos utilizados para construirlas. Registros incompletos, nombres de autores mal normalizados o conceptos ambiguos pueden generar conexiones artificiales o fragmentar relaciones reales. Verma et al. (2023) advierten sobre los desafíos asociados con la estructuración del conocimiento académico, especialmente cuando se integran fuentes heterogéneas. Por ello, antes de interpretar una red es necesario revisar cómo fueron definidas las entidades, qué criterios se usaron para relacionarlas y qué limitaciones presenta el conjunto documental de origen.

Las redes también pueden mostrar diferencias de escala dentro de un mismo campo científico. Algunos autores aparecen vinculados con numerosos colaboradores, ciertos documentos concentran conexiones y determinados conceptos ocupan posiciones recurrentes entre varios grupos. Haunschild y Bornmann (2024) evidencian que los mapas bibliométricos permiten representar la ciencia en niveles individuales, institucionales y nacionales. Esta posibilidad de cambiar la escala facilita observar estructuras amplias y relaciones más específicas, aunque la interpretación debe conservar coherencia entre el nivel analizado y las preguntas que orientan la investigación.

El análisis de software científico ofrece otra vía para comprender cómo se forman redes alrededor de prácticas de investigación. Orduña y Costas (2021), al estudiar el uso de VOSviewer mediante un enfoque basado en enlaces, muestran cómo las relaciones digitales pueden emplearse para seguir la circulación y presencia de herramientas dentro de la comunidad académica. Este tipo de aproximación amplía la noción de red científica, pues incorpora objetos y prácticas que también participan en la producción y organización del conocimiento, además de autores, artículos y conceptos estrictamente considerados.

La lectura de una red requiere distinguir entre centralidad estructural y relevancia sustantiva. Un nodo muy conectado puede ocupar una posición destacada por múltiples razones, y ninguna de ellas garantiza por sí sola que represente el conocimiento más sólido o innovador del campo. Brack et al. (2022) insisten en que los grafos de investigación deben responder a casos de uso concretos y cumplir requisitos de calidad. La interpretación debe preguntarse qué significa cada conexión dentro del contexto analizado antes de atribuir importancia científica a una posición visualmente dominante.

Las redes de conceptos también pueden revelar zonas de contacto entre áreas que utilizan vocabularios diferentes para problemas relacionados. Verma et al. (2023) muestran que los grafos académicos ofrecen mecanismos para estructurar relaciones semánticas dentro de la

comunicación científica. Cuando estas conexiones se integran con redes de documentos y autores, pueden identificarse espacios de convergencia que no siempre aparecen en clasificaciones disciplinares tradicionales. Esta posibilidad resulta valiosa para explorar campos híbridos, siempre que la proximidad conceptual observada sea contrastada con el contenido real de las publicaciones involucradas.

Una red científica adquiere utilidad cuando ayuda a plantear preguntas que una lista de registros no permitiría formular con la misma claridad. Puede mostrar qué autores concentran colaboración, qué documentos conectan conjuntos temáticos o qué conceptos sirven de enlace entre grupos separados. El mapa no resuelve por sí mismo la interpretación, pero sí organiza la complejidad de manera que ciertas relaciones se vuelven visibles. Su función más productiva consiste en orientar una lectura posterior más precisa, donde la estructura observada pueda contrastarse con el contenido y el contexto de la producción científica.

2.3 Análisis de coautoría, cocitación y acoplamiento bibliográfico

El análisis de coautoría, cocitación y acoplamiento bibliográfico permite examinar la producción científica desde relaciones distintas pero complementarias. La coautoría muestra vínculos de colaboración entre investigadores; la cocitación identifica documentos que aparecen citados conjuntamente; y el acoplamiento bibliográfico relaciona publicaciones que comparten referencias. Estas estructuras pueden representarse mediante redes y mapas que facilitan observar agrupamientos, proximidades y conexiones dentro de un campo. Haunschild y Bornmann (2024) muestran que los datos bibliométricos abiertos permiten construir representaciones científicas comparables en distintos niveles de análisis.

Las redes de coautoría permiten estudiar cómo se organiza la colaboración científica alrededor de publicaciones compartidas. Un vínculo entre dos autores indica participación conjunta en uno o más trabajos, mientras la densidad de conexiones puede revelar grupos con colaboración frecuente. Ávila (2024) muestra que los grafos de conocimiento aplicados a datos de investigación permiten representar entidades y relaciones de forma estructurada. Trasladada a la coautoría, esta lógica hace visible la configuración de comunidades académicas, aunque la frecuencia de colaboración no debe interpretarse automáticamente como indicador de calidad científica.

La utilidad del análisis de coautoría aumenta cuando se observa la estructura completa y no únicamente los vínculos individuales. Algunos autores pueden conectar grupos que de otro modo aparecerían separados, mientras otros participan dentro de comunidades más cerradas.

Haunschild y Bornmann (2024) muestran que los mapas bibliométricos pueden construirse para distintos niveles, desde investigadores hasta instituciones y países. Esto permite examinar cómo las relaciones de colaboración cambian según la escala empleada y evita asumir que una estructura visible en un nivel necesariamente se reproduce de la misma forma en otro.

La cocitación responde a una lógica diferente porque no analiza colaboración directa, sino el uso conjunto de documentos dentro de publicaciones posteriores. Cuando dos trabajos son citados repetidamente en los mismos textos, puede existir una relación intelectual o temática entre ellos. Verma et al. (2023) explican que los grafos de conocimiento académico permiten organizar vínculos entre documentos y otros componentes de la comunicación científica. La cocitación puede contribuir así a identificar núcleos de literatura que participan de manera conjunta en la construcción conceptual de un determinado campo de estudio.

Esta relación debe interpretarse con cautela. Dos documentos pueden ser cocitados porque sostienen argumentos semejantes, porque representan enfoques contrapuestos o simplemente porque forman parte de un mismo antecedente bibliográfico. Una conexión frecuente no explica por sí sola el sentido de la relación. Orduña y Costas (2021), al estudiar vínculos asociados con el uso de VOSviewer, muestran la utilidad de analizar relaciones digitales para comprender la presencia de herramientas y prácticas científicas. El principio es aplicable a redes bibliométricas: el enlace necesita una interpretación contextual posterior.

El acoplamiento bibliográfico parte de una relación inversa a la cocitación. En lugar de observar qué publicaciones posteriores citan conjuntamente dos documentos, examina qué referencias comparten dos trabajos analizados. Cuando varias publicaciones recurren a una base bibliográfica semejante, puede inferirse una proximidad temática o intelectual que merece ser examinada. Haunschild y Bornmann (2024) permiten situar este tipo de relaciones dentro de la construcción de mapas bibliométricos, donde la proximidad entre documentos puede utilizarse para reconocer configuraciones amplias de un campo científico.

La diferencia entre cocitación y acoplamiento bibliográfico también afecta la forma de interpretar el tiempo. La cocitación depende de publicaciones posteriores que establezcan relaciones entre documentos previos, mientras el acoplamiento puede calcularse directamente a partir de las referencias compartidas por los trabajos comparados. Verma et al. (2023) muestran que la estructuración del conocimiento científico requiere definir con claridad qué relaciones se representan y cómo se interpretan. Confundir ambos procedimientos puede

conducir a conclusiones erróneas sobre proximidad temática, influencia o evolución de determinadas áreas de investigación.

Figura 3

Comparación de métodos de análisis bibliométrico



Nota. La figura diferencia los principales métodos bibliométricos según el tipo de relación que permiten identificar entre autores, documentos y referencias científicas.

La representación visual de estas relaciones facilita detectar grupos, zonas densamente conectadas y posibles puntos de articulación entre comunidades científicas. Ávila (2024) demuestra que los grafos de conocimiento permiten estructurar y visualizar datos de investigación mediante relaciones explícitas. Sin embargo, la apariencia del mapa depende de decisiones técnicas relativas a selección de datos, criterios de inclusión y tratamiento de vínculos. Un agrupamiento visualmente compacto no constituye una prueba suficiente de homogeneidad intelectual; representa una estructura que debe contrastarse con los documentos que originaron las conexiones.

El análisis bibliométrico apoyado en software puede simplificar la exploración de redes complejas, pero también puede producir una sensación de objetividad excesiva. Orduña y

Costas (2021) muestran la relevancia adquirida por herramientas como VOSviewer dentro de la actividad científica vinculada con análisis de redes. El uso de software facilita procesar grandes conjuntos de relaciones, aunque los resultados siguen dependiendo de decisiones humanas sobre qué datos se incorporan, qué umbrales se aplican y qué tipo de vínculo se considera significativo para responder la pregunta planteada.

La mayor utilidad de estos análisis aparece cuando las redes sirven para orientar una lectura más específica del campo. Una agrupación de autores puede sugerir comunidades de colaboración, una red de cocitación puede señalar referencias que suelen aparecer juntas y el acoplamiento puede revelar conjuntos documentales con bases bibliográficas semejantes. Ninguna de estas estructuras debería convertirse en una interpretación automática del conocimiento. Su valor reside en ofrecer indicios sobre relaciones que después pueden examinarse directamente en las publicaciones, evitando que la representación visual sustituya el análisis conceptual.

2.4 Identificación automática de conceptos y relaciones temáticas

La identificación automática de conceptos busca reconocer unidades de significado dentro de grandes colecciones científicas y establecer relaciones entre ellas. Esta tarea resulta especialmente útil cuando el volumen documental impide revisar manualmente todas las expresiones relevantes de un campo. Jaradeh et al. (2023) muestran que la extracción de información puede alimentar la construcción de grafos de conocimiento mediante la identificación de entidades y vínculos. El interés no se limita a localizar palabras frecuentes, sino a distinguir conceptos con valor analítico y comprender cómo se conectan dentro de la producción científica.

La extracción automática requiere diferenciar entre términos superficiales y conceptos que cumplen una función específica dentro del discurso académico. Un mismo término puede adquirir sentidos diferentes según el contexto, mientras conceptos semejantes pueden aparecer expresados con vocabularios distintos. Verma et al. (2023) explican que los grafos de conocimiento académico dependen de una estructuración adecuada de entidades y relaciones. Esta exigencia convierte la identificación conceptual en un problema de interpretación, no solo de conteo, porque las relaciones científicas rara vez pueden reducirse a coincidencias léxicas simples.

Los modelos de temas ofrecen otra vía para reconocer estructuras conceptuales dentro de corpus amplios. Wu et al. (2024) revisan los modelos neuronales de tópicos y muestran que

estas técnicas permiten descubrir patrones temáticos latentes en colecciones documentales. Su utilidad consiste en agrupar términos y documentos alrededor de configuraciones recurrentes que pueden sugerir áreas de interés. Sin embargo, un tópico generado automáticamente no equivale necesariamente a una categoría científica consolidada. Su interpretación depende de la coherencia semántica de los términos asociados y del contexto documental del que proviene.

El análisis automatizado también puede apoyarse en representaciones distribuidas del lenguaje para identificar proximidades entre documentos y conceptos. Chagnon et al. (2024) evalúan modelos de tópicos aplicados a artículos científicos mediante enfoques basados en representaciones lingüísticas contemporáneas. Este tipo de procedimiento permite captar relaciones que no aparecen mediante frecuencias simples, especialmente cuando existen expresiones diferentes para ideas cercanas. La ventaja analítica radica en ampliar la detección de afinidades temáticas, aunque esas afinidades deben contrastarse con la lectura del contenido antes de convertirlas en categorías interpretativas estables.

Las relaciones temáticas pueden adoptar formas diversas. Algunos conceptos aparecen como componentes de un mismo problema, otros mantienen relaciones jerárquicas y ciertos términos funcionan como puentes entre áreas diferenciadas. Brack et al. (2022) muestran que la construcción de grafos de investigación exige definir con claridad qué tipos de entidades y relaciones serán representados. Esta decisión resulta determinante porque una red temática cambia sustancialmente según se conecten conceptos por coocurrencia, similitud semántica, pertenencia a un mismo documento o relaciones extraídas explícitamente del texto.

Tabla 5

Dimensiones para interpretar la identificación automática de conceptos y relaciones temáticas

Dimensión	Operación automatizada	Resultado esperado	Riesgo interpretativo
Reconocimiento de entidades	Identificar términos o expresiones relevantes	Inventario de conceptos del corpus	Confundir frecuencia con importancia conceptual
Normalización	Agrupar variantes terminológicas	Reducción de duplicidades semánticas	Fusionar conceptos que tienen significados distintos
Detección de tópicos	Agrupar términos y documentos relacionados	Estructuras temáticas latentes	Tratar tópicos estadísticos como categorías teóricas consolidadas
Extracción de relaciones	Identificar vínculos entre entidades	Redes conceptuales estructuradas	Atribuir causalidad a relaciones meramente asociativas

Dimensión	Operación automatizada	Resultado esperado	Riesgo interpretativo
Validación contextual	Contrastar relaciones con los documentos	Interpretación temática más consistente	Aceptar conexiones automáticas sin revisión del contenido

Nota. La tabla diferencia operaciones computacionales y decisiones interpretativas involucradas en la construcción de relaciones conceptuales a partir de literatura científica.

La tabla evidencia que la automatización puede intervenir en diferentes niveles y que cada operación produce un tipo distinto de resultado. Jaradeh et al. (2023) muestran que extraer información para construir grafos requiere procesos sucesivos y no una única clasificación automática. Una entidad detectada debe ser normalizada, relacionada y finalmente interpretada. Esta secuencia es relevante porque los errores acumulados en las primeras etapas pueden alterar la estructura final del mapa, generando conexiones que parecen consistentes visualmente, aunque provengan de identificaciones conceptuales deficientes.

La validación de modelos temáticos constituye, por ello, una etapa necesaria. Vázquez et al. (2022) analizan la validación de modelos de tópicos científicos mediante información estructural y metadatos del corpus, mostrando que la interpretación de agrupamientos puede contrastarse con características externas de los documentos. Esta estrategia evita depender únicamente de la coherencia aparente de las palabras asociadas a cada tema. Si un agrupamiento no mantiene relaciones consistentes con el corpus, su valor para representar una línea científica puede ser limitado, aunque el modelo produzca una separación matemáticamente clara.

Las relaciones conceptuales detectadas automáticamente también pueden contribuir a reconocer puntos de contacto entre áreas que normalmente se examinan por separado. Verma et al. (2023) muestran que los grafos académicos permiten estructurar componentes diversos de la comunicación científica dentro de una misma arquitectura. Cuando los conceptos funcionan como nodos intermedios entre conjuntos documentales, es posible observar conexiones interdisciplinarias que una clasificación tradicional podría ocultar. No obstante, esa proximidad requiere confirmación documental para establecer si existe una relación sustantiva o solo una coincidencia semántica parcial.

La identificación automática resulta más útil cuando se combina con una interpretación orientada por preguntas científicas concretas. Un sistema puede producir decenas de temas y miles de relaciones, pero esa abundancia no garantiza que todas sean relevantes para el análisis.

Wu et al. (2024) muestran que los modelos de tópicos presentan desafíos metodológicos que afectan su interpretación y aplicación. El investigador necesita decidir qué configuraciones merecen atención, cuáles son demasiado generales y qué relaciones pueden aportar una lectura significativa del campo sin convertir la salida algorítmica en una explicación autosuficiente.

Una representación temática adquiere valor cuando permite distinguir estructuras que después pueden ser examinadas con mayor precisión en los documentos originales. La automatización facilita detectar recurrencias, proximidades y conexiones difíciles de observar manualmente, pero no establece por sí sola el significado científico de esos patrones. El criterio decisivo consiste en mantener una separación entre descubrimiento computacional e interpretación conceptual. Los mapas temáticos funcionan mejor cuando orientan preguntas sobre el corpus y no cuando sustituyen la lectura crítica por agrupamientos producidos exclusivamente a partir de similitudes estadísticas.

2.5 Agrupamiento de documentos y formación de clústeres científicos

El agrupamiento automatizado de documentos permite organizar grandes colecciones científicas a partir de semejanzas temáticas, semánticas o estructurales. Su finalidad no consiste únicamente en reducir el volumen de información, sino en identificar conjuntos de publicaciones que comparten patrones suficientemente consistentes para ser analizados como unidades relacionadas. Chagnon et al. (2024) muestran que los modelos de tópicos aplicados a artículos científicos pueden reconocer configuraciones temáticas dentro de corpus extensos. Esta capacidad facilita explorar áreas de concentración documental sin depender exclusivamente de clasificaciones definidas de antemano.

La formación de clústeres parte de una decisión central: establecer qué tipo de semejanza se utilizará para relacionar los documentos. Dos artículos pueden aproximarse porque comparten vocabulario, porque desarrollan temas similares o porque mantienen relaciones dentro de una estructura científica más amplia. Wu et al. (2024) muestran que los modelos neuronales de tópicos permiten representar patrones latentes en colecciones textuales, aunque sus resultados dependen del enfoque empleado. Por ello, un clúster debe interpretarse según el criterio que originó su formación y no como una categoría universal del campo.

Los modelos temáticos resultan útiles cuando el corpus contiene cientos o miles de publicaciones y la revisión manual dificulta reconocer configuraciones generales. Chagnon et al. (2024) analizan procedimientos para comparar modelos de tópicos sobre artículos científicos, evidenciando que distintas representaciones pueden producir agrupamientos

diferentes. Esta variación tiene consecuencias metodológicas, porque la estructura encontrada no constituye una propiedad completamente independiente del modelo utilizado. La selección del procedimiento de agrupamiento debe responder al tipo de relación que se pretende estudiar y al nivel de detalle requerido.

Otra posibilidad consiste en combinar similitud temática con información procedente de redes y metadatos científicos. Vázquez et al. (2022) proponen validar modelos de tópicos mediante análisis de grafos y características del corpus, lo que permite contrastar si los agrupamientos obtenidos mantienen relaciones coherentes con otras propiedades documentales. Esta integración resulta relevante porque un conjunto puede parecer temáticamente consistente a partir del lenguaje y, sin embargo, presentar poca correspondencia con estructuras de citación, categorías disciplinares u otros elementos que ayuden a interpretar su significado dentro del campo.

Los clústeres pueden facilitar la identificación de núcleos de investigación relativamente consolidados, zonas de contacto entre temas y conjuntos documentales con menor conexión respecto del resto del corpus. Haunschild y Bornmann (2024) muestran que los mapas bibliométricos permiten representar configuraciones científicas a distintas escalas mediante datos abiertos. Cuando estas representaciones incorporan agrupamientos, el investigador puede examinar la distribución del conocimiento de manera más estructurada. Sin embargo, la proximidad visual entre documentos debe entenderse como resultado de criterios de cálculo específicos y no como prueba suficiente de identidad conceptual.

Tabla 6

Dimensiones para analizar la formación de clústeres en colecciones científicas

Dimensión	Base del agrupamiento	Aporte analítico	Riesgo de interpretación
Similitud semántica	Proximidad de significados entre documentos	Detectar publicaciones temáticamente cercanas	Agrupar textos con vocabulario similar, pero propósitos distintos
Tópicos latentes	Patrones recurrentes de términos	Reconocer estructuras temáticas no predefinidas	Convertir un patrón estadístico en categoría conceptual rígida
Relaciones de red	Vínculos entre documentos o entidades	Identificar comunidades estructurales	Confundir conectividad con afinidad intelectual
Metadatos	Año, disciplina, autoría u otras características	Contrastar la coherencia del agrupamiento	Sobrevalorar atributos externos frente al contenido

Dimensión	Base del agrupamiento	Aporte analítico	Riesgo de interpretación
Validación del clúster	Comparación entre estructura y corpus	Evaluar consistencia e interpretabilidad	Aceptar agrupamientos solo por su apariencia visual

Nota. La tabla organiza criterios que permiten distinguir cómo se forman y cómo deben interpretarse los clústeres científicos.

La comparación entre estas dimensiones muestra que no existe una única forma válida de agrupar literatura científica. Vázquez et al. (2022) evidencian que la validación puede enriquecerse cuando la estructura temática se contrasta con información adicional del corpus. Esta práctica evita interpretar un agrupamiento únicamente por las palabras predominantes que contiene. Un clúster adquiere mayor valor analítico cuando existe coherencia entre su composición textual, las relaciones entre documentos y otras características pertinentes, aunque ninguna de estas señales debería utilizarse de manera aislada para definir una línea científica.

Los grafos de conocimiento ofrecen otra posibilidad para organizar agrupamientos cuando las relaciones entre documentos se combinan con entidades científicas adicionales. Verma et al. (2023) muestran que estas estructuras pueden integrar publicaciones, autores, conceptos y otros componentes de la comunicación académica. Un clúster construido dentro de este entorno puede reflejar más que semejanza textual, porque incorpora conexiones procedentes de diferentes tipos de relaciones. Esta riqueza también aumenta la dificultad interpretativa, ya que la pertenencia de un documento al grupo puede responder a varios criterios simultáneamente.

La estabilidad de los clústeres merece atención cuando se pretende utilizarlos para describir un campo. Wu et al. (2024) muestran que los modelos de tópicos presentan desafíos relacionados con su evaluación e interpretación, lo que obliga a examinar si los agrupamientos se mantienen ante cambios razonables en el procedimiento. Una estructura excesivamente sensible a pequeñas modificaciones puede ofrecer una imagen poco robusta del corpus. La consistencia no exige que todos los métodos produzcan exactamente los mismos grupos, pero sí que las principales relaciones puedan justificarse con evidencia documental suficiente.

Un clúster científico resulta útil cuando permite formular una hipótesis de organización que después pueda contrastarse con los documentos que lo componen. Su valor no reside en asignar etiquetas automáticas y asumirlas como descripciones definitivas del campo. Los agrupamientos pueden revelar afinidades, fronteras difusas o conexiones inesperadas, pero cada una de esas configuraciones requiere interpretación. Cuando el mapa se utiliza como punto de partida y no como respuesta final, la agrupación automática contribuye a explorar la

estructura del conocimiento sin convertir una salida computacional en una clasificación incuestionable.

2.6 Visualización interactiva de campos y líneas de investigación

La visualización interactiva permite representar campos científicos complejos mediante mapas que organizan documentos, autores, conceptos y relaciones en estructuras navegables. Su utilidad radica en convertir grandes volúmenes de información en configuraciones que pueden explorarse desde distintos niveles de detalle. Ávila (2024) muestra que los grafos de conocimiento aplicados a datos abiertos de investigación facilitan la representación de vínculos entre entidades científicas. Esta posibilidad ayuda a observar patrones de concentración, conexiones temáticas y zonas periféricas que serían difíciles de identificar mediante una revisión exclusivamente lineal de publicaciones.

La interacción modifica la forma de interpretar estos mapas porque permite desplazarse entre escalas, seleccionar nodos y examinar relaciones específicas sin perder la visión general del campo. Haunschild y Bornmann (2024) muestran que OpenAlex puede utilizarse para construir mapas bibliométricos capaces de representar la ciencia en niveles individuales, institucionales y nacionales. Esta flexibilidad permite adaptar la lectura a la pregunta de investigación, aunque exige mantener claridad sobre qué unidad se está observando. Una relación significativa en una escala puede perder relevancia cuando se analiza desde otra.

Las líneas de investigación pueden hacerse visibles cuando ciertos grupos de documentos, conceptos o autores muestran una concentración estable de relaciones. Orduña y Costas (2021), al estudiar el uso de VOSviewer, muestran la importancia que han adquirido las herramientas de visualización en el análisis de estructuras científicas. Estas representaciones ayudan a identificar agrupamientos y proximidades, pero no deben confundirse con una clasificación definitiva del conocimiento. Una línea de investigación puede presentar fronteras difusas, compartir conceptos con otros grupos o cambiar conforme se incorporan nuevas publicaciones al corpus analizado.

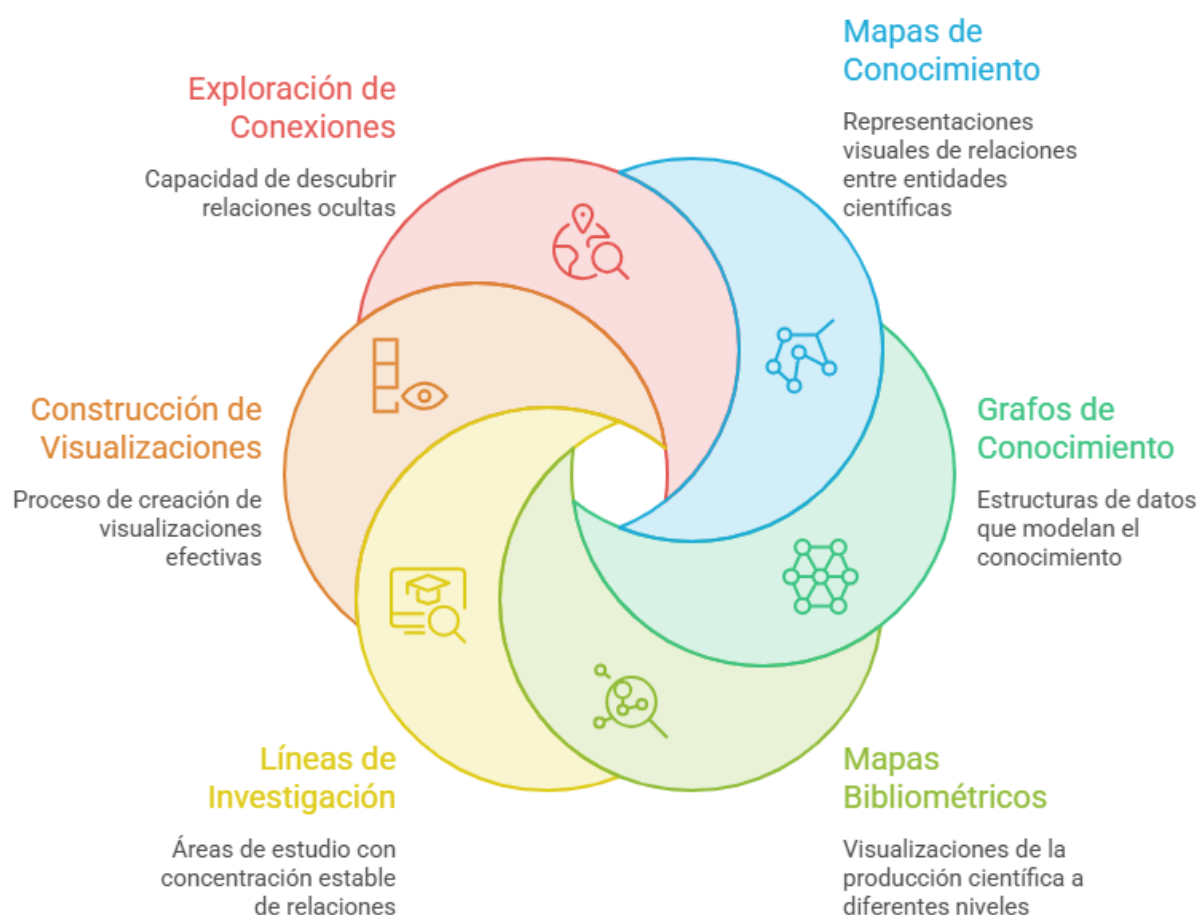
La construcción de una visualización depende de los datos que alimentan el sistema y de las relaciones seleccionadas para representarlos. Brack et al. (2022) destacan que un grafo abierto de conocimiento de investigación necesita atender tanto a los casos de uso como a requisitos de calidad y estrategias de construcción. Esta condición resulta decisiva cuando se pretende interpretar campos científicos, porque una visualización puede parecer completa, aunque

proceda de datos parciales. Las ausencias documentales, errores de metadatos o relaciones mal definidas pueden alterar la imagen final del campo.

La interacción también permite explorar conexiones que no resultan evidentes en una representación estática. Un usuario puede aislar un conjunto temático, seguir vínculos entre autores o examinar cómo determinados documentos conectan grupos diferentes. Vázquez et al. (2022) muestran que el análisis de grafos y los metadatos del corpus pueden contribuir a validar estructuras temáticas obtenidas mediante modelos de tópicos. Esta combinación ofrece una vía para interpretar mapas no solo por su apariencia visual, sino también por la coherencia entre agrupamientos, relaciones documentales y características del corpus.

Figura 4

Explorando el paisaje científico con visualización interactiva



Nota. La figura integra distintas formas de visualización utilizadas para representar relaciones, estructuras y áreas de concentración dentro de la producción científica.

La representación de campos científicos mediante mapas superpuestos ofrece otra posibilidad analítica. Haunschild y Bornmann (2024) plantean el uso de mapas bibliométricos globales para situar unidades de análisis dentro de una estructura científica más amplia. Esta estrategia

permite comparar posiciones relativas y observar hacia qué áreas se orienta una producción determinada. Sin embargo, la interpretación depende de la base utilizada para construir el mapa de referencia. Una estructura general demasiado agregada puede ocultar diferencias internas importantes entre especialidades, subcampos o líneas emergentes de investigación.

Las visualizaciones interactivas pueden facilitar también la detección de zonas de transición entre áreas. Cuando varios grupos se conectan mediante conceptos, documentos o autores compartidos, el mapa puede mostrar espacios donde convergen tradiciones o problemas diferentes. Ávila (2024) evidencia que los grafos de conocimiento permiten integrar relaciones entre datos de investigación, lo que favorece una lectura menos fragmentada del campo. Estas zonas intermedias merecen especial atención porque pueden representar vínculos sustantivos, aunque también pueden ser producto de relaciones generales que aparecen en numerosos contextos científicos.

El uso de software especializado simplifica el procesamiento de redes extensas, pero introduce decisiones técnicas que condicionan la visualización. Orduña y Costas (2021) muestran la presencia de VOSviewer en prácticas de análisis científico, lo que evidencia la consolidación de herramientas diseñadas para representar relaciones bibliométricas. La disposición espacial, los umbrales de inclusión y los criterios de agrupamiento influyen en la forma final del mapa. Por ello, dos visualizaciones construidas con el mismo corpus pueden destacar aspectos diferentes si utilizan configuraciones distintas.

Una visualización interactiva también puede utilizarse para estudiar cambios en las líneas de investigación cuando incorpora información temporal. Aunque el mapa estático muestra relaciones en un momento determinado, la organización por periodos permite observar desplazamientos en la concentración temática o en la posición de ciertos grupos. Vázquez et al. (2022) aportan elementos para relacionar estructuras temáticas con metadatos del corpus, lo que facilita examinar la distribución de temas bajo distintos criterios. Esta lectura temporal requiere evitar interpretaciones causales cuando los datos solo muestran variaciones de presencia o conexión.

La interpretación crítica de estos mapas exige reconocer que la prominencia visual no equivale necesariamente a importancia científica. Un nodo grande o central puede reflejar frecuencia, conectividad o criterios definidos por el software, pero esas propiedades no sustituyen la evaluación del contenido. Brack et al. (2022) insisten en la necesidad de construir grafos adecuados a objetivos concretos y con requisitos explícitos de calidad. La lectura responsable

debe identificar qué representa cada elemento antes de atribuirle centralidad conceptual, influencia o relevancia dentro de un campo determinado.

La visualización interactiva resulta más productiva cuando funciona como instrumento de exploración y no como sustituto de la lectura científica. Un mapa puede orientar hacia agrupamientos, conexiones o zonas poco desarrolladas, pero esas señales adquieren significado solo cuando se contrastan con los documentos que las originan. La posibilidad de navegar entre escalas y relaciones permite formular preguntas más precisas sobre la estructura del campo. Su valor analítico depende de conservar esa apertura interpretativa y evitar que la claridad gráfica convierta relaciones complejas en conclusiones demasiado simples.

2.7 Interpretación crítica de mapas científicos generados con IA

Los mapas científicos generados con inteligencia artificial ofrecen una representación sintética de relaciones complejas entre documentos, conceptos, autores y áreas de conocimiento. Su utilidad depende de que sean interpretados como construcciones analíticas y no como reproducciones directas de la realidad científica. Vázquez et al. (2022) muestran que las estructuras temáticas derivadas de modelos pueden validarse mediante grafos y metadatos del corpus, lo que obliga a examinar la coherencia entre la configuración visual y las propiedades documentales que la sustentan antes de formular conclusiones sobre un campo.

La disposición espacial de nodos y clústeres puede inducir interpretaciones intuitivas que no siempre corresponden con el significado científico de las relaciones representadas. Haunschild y Bornmann (2024) muestran que los mapas bibliométricos permiten situar unidades dentro de estructuras científicas amplias, pero su lectura exige comprender qué indicadores y criterios determinan la posición de cada elemento. La cercanía visual puede expresar similitud, conexión o pertenencia según el método aplicado. Sin esa información, existe el riesgo de atribuir significado sustantivo a una proximidad producida por decisiones técnicas.

La interpretación crítica comienza, por tanto, con una revisión del proceso de construcción del mapa. Debe conocerse qué corpus fue utilizado, qué relaciones se modelaron y qué información quedó fuera de la representación. Brack et al. (2022) señalan que la calidad de un grafo de conocimiento depende de sus datos, requisitos y estrategias de construcción. Un mapa generado a partir de información incompleta puede mostrar una estructura internamente coherente y, al mismo tiempo, ofrecer una imagen parcial del campo. La apariencia de orden no compensa las limitaciones de cobertura.

La inteligencia artificial puede facilitar la detección de agrupamientos y relaciones que serían difíciles de reconocer manualmente, pero también introduce opacidad cuando el usuario desconoce los criterios empleados por el modelo. Zárte (2022) aborda la explicabilidad como un problema central de los sistemas de inteligencia artificial. Aplicado a los mapas científicos, este aspecto implica poder comprender por qué determinados documentos aparecen conectados, por qué ciertos conceptos se agrupan y qué variables intervienen en su posición relativa. Una visualización difícil de explicar limita su valor como instrumento científico.

La interpretabilidad resulta especialmente importante cuando el mapa influye en decisiones posteriores, como seleccionar líneas de investigación, identificar núcleos temáticos o definir documentos prioritarios para lectura. Ciobanu et al. (2024) destacan la importancia de sistemas reproducibles e interpretables en aplicaciones de aprendizaje automático. Aunque su trabajo se desarrolla en medicina, el principio es pertinente para el análisis científico: una representación útil debe permitir examinar el proceso que condujo a sus resultados. Sin esa posibilidad, el mapa puede orientar decisiones sin ofrecer elementos suficientes para evaluar su confiabilidad.

Otra dificultad proviene de las desigualdades existentes en la producción y circulación del conocimiento. Ross et al. (2022) analizan dinámicas de ventaja acumulativa y amenazas a la equidad en la ciencia abierta, mostrando que la visibilidad científica no se distribuye de forma uniforme. Si un mapa se construye a partir de bases donde ciertos autores, instituciones o regiones poseen mayor presencia, esas desigualdades pueden reproducirse en la representación. Un nodo central puede reflejar tanto actividad científica como condiciones previas de visibilidad y acceso dentro del sistema documental utilizado.

La centralidad debe interpretarse, por ello, con especial cautela. Un autor o documento muy conectado no necesariamente constituye el aporte conceptualmente más importante del campo. Haunschild y Bornmann (2024) muestran que los mapas bibliométricos permiten observar posiciones relativas dentro de estructuras científicas, pero estas posiciones responden a métricas concretas. El análisis crítico debe diferenciar entre frecuencia, conectividad, citación y relevancia sustantiva. Confundir estas dimensiones puede convertir propiedades estructurales de la red en juicios de valor que la representación no está en condiciones de demostrar.

Los clústeres también requieren una lectura cuidadosa. Vázquez et al. (2022) evidencian que los modelos de tópicos necesitan procedimientos de validación que permitan comprobar si los agrupamientos guardan coherencia con otras características del corpus. Una agrupación visualmente definida puede contener documentos con diferencias importantes de enfoque,

método o contexto. La etiqueta asignada al clúster funciona como una interpretación resumida de un conjunto heterogéneo y no como una descripción exhaustiva. Revisar documentos representativos ayuda a determinar hasta qué punto el nombre atribuido expresa realmente su contenido.

La posibilidad de comparar mapas generados con distintos parámetros puede convertirse en una estrategia de control. Ciobanu et al. (2024) relacionan reproducibilidad e interpretabilidad con la necesidad de examinar críticamente los resultados obtenidos mediante aprendizaje automático. Si pequeñas modificaciones en los criterios de agrupamiento producen estructuras radicalmente diferentes, la estabilidad de la interpretación merece ser cuestionada. La comparación entre configuraciones no busca encontrar un único mapa verdadero, sino identificar qué relaciones persisten y cuáles dependen en mayor medida de decisiones técnicas específicas.

Un mapa científico debe considerarse una hipótesis visual sobre la organización de un corpus. Su principal aporte consiste en señalar relaciones que merecen ser investigadas, no en establecer por sí mismo cuáles son las fronteras definitivas de un campo. La lectura crítica exige volver a los documentos, revisar la procedencia de los datos y comprobar si los patrones visuales resisten una interpretación conceptual más profunda. Cuando se mantiene esa distancia analítica, la inteligencia artificial amplía la capacidad de exploración sin convertir la representación gráfica en una explicación autosuficiente del conocimiento científico.

Capítulo

3.

Lectura automatizada de grandes colecciones de artículos

Evelin Mariela Miranda Hermosa



3.1 Procesamiento automatizado del contenido científico

El procesamiento automatizado del contenido científico busca transformar grandes volúmenes de publicaciones en información susceptible de ser organizada, clasificada y analizada mediante procedimientos computacionales. Esta transformación resulta especialmente útil cuando la cantidad de documentos supera la capacidad de revisión manual disponible. Ofori et al. (2024) examinan la aplicación del procesamiento del lenguaje natural, el aprendizaje automático y el aprendizaje profundo en revisiones sistemáticas, mostrando que estas técnicas pueden intervenir en distintas fases del tratamiento documental. Su utilidad depende de vincular la automatización con objetivos de análisis previamente definidos.

Procesar un artículo científico implica trabajar con una estructura documental compleja en la que título, resumen, metodología, resultados y referencias cumplen funciones diferentes. Los sistemas automatizados deben distinguir qué información resulta relevante según la tarea que se pretende ejecutar. Torre et al. (2023) muestran que la inteligencia artificial puede apoyar la automatización de revisiones de literatura mediante procedimientos aplicados a diferentes etapas del proceso. Esta diferenciación resulta necesaria porque identificar un documento pertinente no exige las mismas operaciones que extraer información específica de su contenido o sintetizar sus principales aportes.

El procesamiento del lenguaje natural permite analizar contenido textual mediante procedimientos capaces de reconocer patrones, relaciones y unidades informativas. Hua et al. (2023) describen avances en algoritmos inteligentes aplicados al cálculo científico, situando estas técnicas dentro de un marco más amplio de automatización de tareas complejas. En el análisis documental, su aplicación puede facilitar la transformación de textos extensos en representaciones manejables. No obstante, la capacidad de procesar información rápidamente no implica que el sistema comprenda necesariamente el alcance metodológico o conceptual de cada fragmento analizado.

Una de las ventajas operativas del procesamiento automatizado es la posibilidad de aplicar criterios semejantes a grandes conjuntos documentales. Esta regularidad puede reducir parte de la variabilidad asociada con revisiones manuales extensas, especialmente durante tareas repetitivas. Ofori et al. (2024) muestran que técnicas automatizadas pueden apoyar procedimientos de selección, clasificación y extracción dentro de revisiones sistemáticas. Sin embargo, la uniformidad computacional no debe confundirse con neutralidad, porque los

resultados dependen de los modelos utilizados, de los datos procesados y de las reglas establecidas para interpretar cada documento.

La organización previa del corpus influye directamente en la calidad del procesamiento posterior. Tapia y Recalde (2024) destacan la importancia de estructurar adecuadamente los flujos de gestión documental para facilitar el tratamiento sistemático de la información. En colecciones científicas extensas, esta exigencia supone controlar duplicados, formatos, metadatos y versiones antes de aplicar procedimientos automatizados. Un corpus desordenado puede producir resultados inconsistentes incluso cuando el algoritmo empleado funciona correctamente. La calidad del análisis comienza, por tanto, con una preparación documental que reduzca errores evitables en las etapas posteriores.

El procesamiento automatizado también puede utilizarse para reducir la cantidad de información que requiere lectura detallada. Torre et al. (2023) muestran que la inteligencia artificial puede incorporarse a procesos estructurados de revisión, facilitando el tratamiento de grandes conjuntos de publicaciones. Esta capacidad permite concentrar el esfuerzo humano en documentos o fragmentos con mayor probabilidad de aportar información relevante. El ahorro de tiempo, sin embargo, depende de que los criterios de automatización estén bien definidos; una clasificación deficiente puede desplazar documentos valiosos hacia zonas poco visibles del corpus analizado.

La síntesis automatizada representa una extensión natural de estas capacidades, pues convierte información distribuida en textos más condensados. Tang et al. (2023) evalúan modelos de lenguaje aplicados a la síntesis de evidencia médica y muestran que estas herramientas pueden procesar información científica para producir representaciones resumidas. La utilidad de esta función aumenta cuando se manejan colecciones numerosas, aunque también introduce riesgos asociados con omisiones o simplificaciones. Una síntesis automática debe considerarse una representación derivada del contenido original y no un sustituto suficiente de la revisión crítica de la fuente.

Los modelos de lenguaje pueden desempeñar funciones de curación de información cuando se integran con mecanismos destinados a reducir respuestas no sustentadas. Gilbert et al. (2024) examinan modelos aumentados orientados a la curación de información médica sin alucinaciones, destacando la importancia de vincular la generación con fuentes verificables. Este enfoque resulta pertinente para el procesamiento científico porque permite distinguir entre reorganizar evidencia existente y producir contenido no comprobado. La calidad del sistema

depende de mantener una relación trazable entre la salida generada y los documentos que sirvieron como base.

El procesamiento automatizado adquiere mayor valor cuando conserva la posibilidad de reconstruir qué operaciones fueron aplicadas a cada documento. Kapoor y Narayanan (2023) advierten sobre problemas de reproducibilidad en ciencia basada en aprendizaje automático, especialmente cuando los procedimientos no permiten identificar adecuadamente cómo se produjeron los resultados. En colecciones científicas, la trazabilidad debe incluir criterios de selección, transformaciones textuales y decisiones de clasificación. Sin ese registro, resulta difícil distinguir si una conclusión proviene del contenido original o de una operación computacional que modificó su representación durante el análisis.

La lectura automatizada no elimina la necesidad de decidir qué partes del conocimiento científico pueden ser procesadas de forma fiable y cuáles requieren interpretación especializada. Un sistema puede clasificar fragmentos, identificar patrones o condensar información, pero estas operaciones tienen límites cuando el sentido depende de matices metodológicos, supuestos teóricos o relaciones argumentativas complejas. El diseño adecuado del flujo de trabajo debe separar tareas repetitivas de tareas interpretativas. Esa distinción evita convertir la automatización en una sustitución indiscriminada de la lectura y permite utilizarla como apoyo para ampliar la capacidad de análisis.

3.2 Extracción de información relevante de textos académicos

La extracción de información relevante permite localizar dentro de un artículo aquellos elementos que responden de manera directa a una necesidad de análisis. En colecciones extensas, esta tarea puede orientarse a identificar objetivos, métodos, variables, poblaciones, resultados, conceptos o relaciones específicas. Jaradeh et al. (2023) muestran que los procesos de extracción de información pueden transformar contenido textual no estructurado en entidades y vínculos utilizables posteriormente. Esta operación reduce la necesidad de revisar de forma lineal cada documento, aunque exige definir con precisión qué tipo de información se considera relevante.

La selección de información depende del propósito de la revisión. Un mismo artículo puede contener datos metodológicos útiles para una pregunta y resultados conceptuales relevantes para otra. Ofori et al. (2024) muestran que el procesamiento del lenguaje natural y otras técnicas de inteligencia artificial pueden apoyar tareas de extracción dentro de revisiones sistemáticas. Esta capacidad no elimina la necesidad de establecer categorías previas de análisis. Si el

sistema desconoce qué variables, atributos o relaciones deben recuperarse, puede producir una salida abundante pero poco útil para construir evidencia comparable entre documentos.

La extracción automatizada también debe reconocer que los textos académicos presentan estructuras heterogéneas. Aunque muchos artículos organizan su contenido en secciones relativamente estables, la ubicación y forma de reportar información puede variar entre disciplinas y tipos de estudio. Torre et al. (2023) muestran que la automatización de revisiones requiere adaptar procedimientos a diferentes etapas del proceso documental. En este contexto, localizar una muestra, un resultado o una limitación no consiste solo en encontrar determinadas palabras, sino en interpretar la función que esas expresiones cumplen dentro de la estructura del artículo.

Los grafos de conocimiento ofrecen una forma de convertir la información extraída en relaciones estructuradas. Verma et al. (2023) explican que estas representaciones pueden integrar documentos, autores, conceptos y otras entidades dentro de un mismo sistema. La extracción deja entonces de producir únicamente listas de datos y puede generar vínculos entre elementos identificados en múltiples publicaciones. Esta transformación resulta útil cuando se pretende comparar cómo un concepto se relaciona con determinados métodos, poblaciones o resultados a lo largo de un corpus amplio, siempre que las relaciones extraídas sean suficientemente precisas.

Tabla 7

Dimensiones para la extracción automatizada de información en textos académicos

Dimensión	Información buscada	Utilidad analítica	Riesgo principal
Objetivo del estudio	Propósito o pregunta investigada	Comparar finalidades entre documentos	Confundir antecedentes con objetivos explícitos
Metodología	Diseño, muestra, técnicas y procedimientos	Identificar patrones metodológicos	Extraer datos aislados sin contexto
Resultados	Hallazgos y relaciones reportadas	Contrastar evidencia entre estudios	Separar resultados de sus condiciones de interpretación
Conceptos	Términos, categorías y entidades relevantes	Construir relaciones temáticas	Tratar palabras frecuentes como conceptos equivalentes
Limitaciones	Restricciones reconocidas en el estudio	Valorar alcance de los hallazgos	Omitir limitaciones expresadas de forma indirecta

Nota. La tabla organiza tipos de información que pueden extraerse de forma automatizada y los principales problemas que requieren comprobación durante su uso.

La extracción de resultados requiere especial cuidado porque una afirmación científica suele depender de condiciones metodológicas que no pueden separarse sin pérdida de sentido. Tang et al. (2023), al evaluar modelos de lenguaje para sintetizar evidencia médica, muestran la complejidad de procesar información científica de manera condensada. Un resultado extraído sin información sobre población, comparación o procedimiento puede adquirir una apariencia de generalidad que el estudio original no sostiene. Por ello, los sistemas deberían conservar suficiente contexto para interpretar correctamente cada dato recuperado.

Otro problema aparece cuando una misma entidad se expresa de formas diferentes a lo largo del corpus. Jaradeh et al. (2023) muestran que la construcción de grafos a partir de información extraída requiere procedimientos de reconocimiento y estructuración capaces de manejar entidades y relaciones. La normalización resulta necesaria para evitar que variantes terminológicas sean tratadas como conceptos independientes. Sin embargo, agrupar automáticamente expresiones semejantes también puede borrar diferencias importantes, especialmente cuando términos próximos adquieren significados particulares dentro de disciplinas, diseños o marcos conceptuales diferentes.

La gestión del corpus condiciona igualmente la calidad de la extracción. Tapia y Recalde (2024) permiten situar la organización documental como un componente previo del tratamiento sistemático de información. Si existen registros duplicados, archivos incompletos o documentos mal identificados, los resultados extraídos pueden incorporar redundancias o errores difíciles de detectar posteriormente. Una colección bien estructurada permite relacionar cada dato con su fuente original y facilita revisar discrepancias. La trazabilidad debe conservarse incluso cuando la extracción se realiza de manera masiva mediante herramientas automatizadas.

La información extraída tampoco debe interpretarse como una representación exhaustiva del artículo. Todo procedimiento selecciona ciertos elementos y deja otros fuera según las categorías utilizadas. Ofori et al. (2024) muestran que la automatización de revisiones puede apoyar tareas específicas, pero su utilidad depende del diseño del flujo de trabajo. Una matriz de datos puede facilitar comparaciones entre cientos de documentos y, al mismo tiempo, perder matices argumentativos que solo aparecen mediante lectura contextual. La extracción debe servir para orientar el análisis, no para reducir cada publicación a una secuencia rígida de campos.

El criterio más exigente consiste en conservar la relación entre el dato extraído y el fragmento documental que lo sustenta. Cuando esa conexión se pierde, la información puede circular dentro del análisis sin que sea posible comprobar su origen o contexto. Un sistema útil debería permitir regresar desde una categoría, entidad o resultado hacia la publicación correspondiente y revisar cómo fue expresado. Esta trazabilidad convierte la extracción automatizada en una herramienta verificable y evita que la eficiencia del procesamiento termine produciendo datos aparentemente precisos pero desvinculados de la evidencia original.

3.3 Resumen automático de artículos y documentos científicos

El resumen automático de artículos científicos busca condensar textos extensos en representaciones más breves que conserven sus ideas centrales, resultados principales y relaciones relevantes. Esta capacidad resulta especialmente útil cuando el investigador trabaja con colecciones documentales amplias y necesita una primera aproximación al contenido antes de realizar una lectura detallada. Tang et al. (2023) evalúan modelos de lenguaje aplicados a la síntesis de evidencia médica y muestran que estas herramientas pueden generar resúmenes útiles, aunque su calidad depende de la fidelidad con que preserven la información contenida en las fuentes originales.

La síntesis automatizada no consiste únicamente en reducir la longitud de un texto. Un buen resumen debe conservar la estructura argumentativa suficiente para que el lector comprenda qué problema se estudió, cómo se abordó y qué resultados fueron reportados. Gao et al. (2023), al comparar resúmenes científicos generados por ChatGPT con resúmenes reales, muestran que los modelos pueden producir textos con apariencia académica convincente. Esta capacidad obliga a diferenciar entre fluidez lingüística y precisión científica, porque una redacción coherente puede contener omisiones o formulaciones que no reflejen adecuadamente el documento de origen.

La selección de contenido constituye uno de los principales desafíos. Cuando un modelo decide qué información conservar, puede priorizar ciertos hallazgos y dejar fuera condiciones metodológicas necesarias para interpretarlos. Tang et al. (2023) muestran que la evaluación de resúmenes de evidencia debe considerar no solo legibilidad, sino también fidelidad y utilidad informativa. En textos científicos, eliminar detalles sobre población, diseño o condiciones experimentales puede modificar el alcance de una conclusión. La brevedad, por tanto, debe equilibrarse con la conservación de elementos indispensables para comprender correctamente los resultados.

La generación automática también puede introducir información que no aparece de forma explícita en el documento original. Gilbert et al. (2024) analizan modelos de lenguaje aumentados orientados a reducir alucinaciones durante la curación de información médica, destacando la importancia de vincular las respuestas con fuentes verificables. Este problema es particularmente relevante en los resúmenes científicos, porque una frase añadida por inferencia puede parecer una conclusión del estudio. La síntesis debe distinguir con claridad entre contenido extraído, reformulación y cualquier interpretación que exceda lo expresado por la fuente.

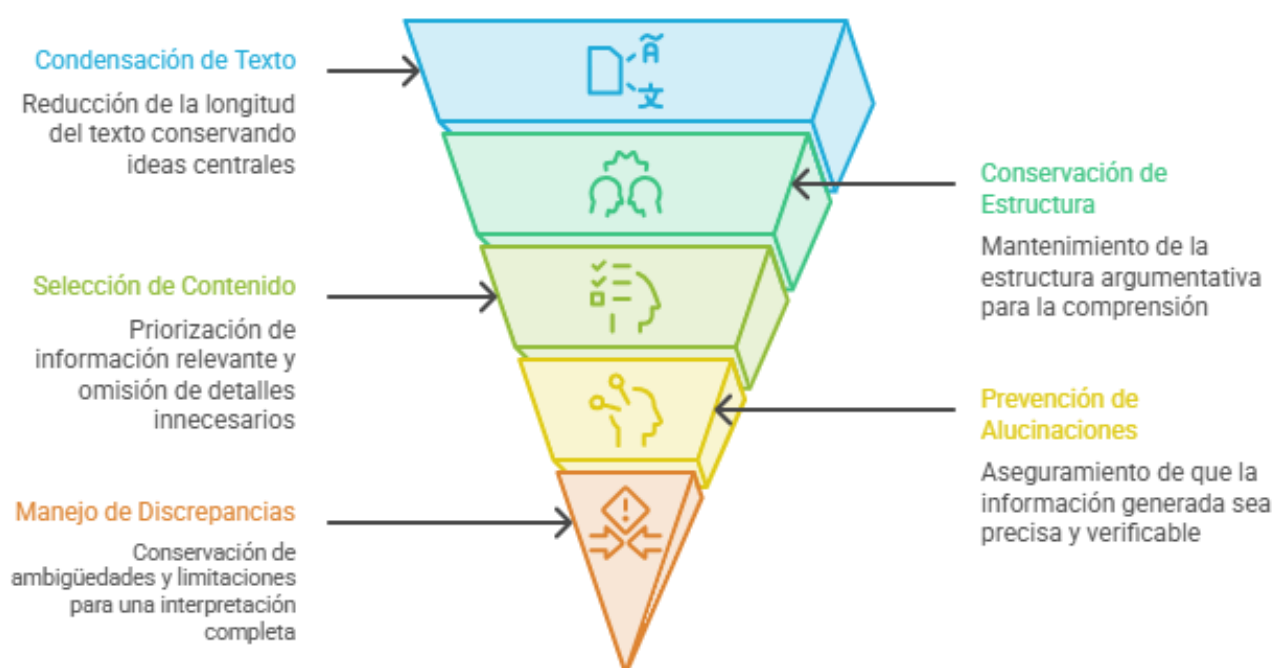
La extensión del documento influye en la dificultad del procesamiento. Un artículo breve y estructurado ofrece condiciones diferentes a una revisión extensa o a un conjunto de publicaciones combinadas. Krenn et al. (2022) analizan las posibilidades de la inteligencia artificial para contribuir a la comprensión científica y señalan que estos sistemas pueden apoyar tareas complejas de procesamiento de conocimiento. Sin embargo, la comprensión científica implica relaciones conceptuales que no siempre pueden reducirse a una selección de frases. Un resumen útil debe conservar conexiones entre argumentos y no limitarse a enumerar resultados aislados.

Los resúmenes automáticos pueden emplearse como herramientas de cribado preliminar. Cuando el corpus contiene numerosos documentos, una síntesis inicial permite decidir cuáles requieren lectura completa y cuáles presentan menor relación con el objetivo de investigación. Doneva et al. (2024) examinan el uso de grandes modelos de lenguaje para procesar, analizar y sintetizar textos biomédicos, mostrando que estas tecnologías pueden apoyar distintas tareas documentales. Su empleo resulta más adecuado cuando el resumen funciona como una capa intermedia de exploración y no como sustituto definitivo de la lectura del artículo original.

Otra cuestión crítica es la posibilidad de que el resumen elimine discrepancias internas del documento. Un artículo puede presentar resultados ambiguos, limitaciones importantes o diferencias entre la interpretación de los autores y los datos reportados. Gao et al. (2023) muestran que los modelos generativos pueden producir textos académicamente verosímiles, lo que incrementa el riesgo de que una síntesis demasiado uniforme suavice tensiones presentes en la fuente. Cuando el objetivo es comparar evidencia científica, estas diferencias deben conservarse porque pueden afectar directamente la interpretación posterior del conjunto documental.

La síntesis automática adquiere mayor valor cuando permite rastrear cada afirmación hasta el fragmento correspondiente del documento. Gilbert et al. (2024) muestran la importancia de sistemas aumentados que relacionen la generación con información recuperada de fuentes verificables. En el contexto académico, esta trazabilidad permite comprobar si una frase del resumen representa fielmente el contenido original. Además, facilita revisar rápidamente aquellas partes que parecen ambiguas o especialmente relevantes, reduciendo la dependencia de una salida generada cuya precisión no puede asumirse sin verificación documental.

Figura 5
Proceso de resumen automático de artículos científicos



Nota. La figura representa los criterios que intervienen en la generación de resúmenes automáticos orientados a conservar la información esencial y reducir distorsiones del contenido científico.

El nivel de detalle también debe adaptarse a la finalidad del resumen. Una síntesis destinada a decidir pertinencia puede concentrarse en objetivo y alcance, mientras una utilizada para comparar resultados requiere información metodológica y empírica más precisa. Tang et al. (2023) evidencian que la utilidad de los modelos depende del tipo de evidencia que deben resumir y del criterio con que se evalúa la salida. Esto implica que no existe un único formato adecuado para toda tarea científica; el contenido seleccionado debe responder a la función que tendrá dentro del proceso de análisis.

La automatización puede generar una falsa sensación de comprensión cuando permite recorrer rápidamente decenas de artículos mediante resúmenes breves. Krenn et al. (2022) plantean que

la inteligencia artificial puede contribuir al entendimiento científico, pero esa contribución exige reconocer las diferencias entre procesar información y comprender sus implicaciones. Un resumen puede orientar hacia los elementos principales de un estudio, aunque no siempre refleja supuestos teóricos, decisiones metodológicas o limitaciones implícitas. La lectura completa sigue siendo necesaria cuando esas dimensiones afectan la interpretación de la evidencia.

El resumen automático resulta más valioso cuando funciona como una interfaz entre la abundancia documental y la lectura crítica. Su propósito no debería ser convertir artículos complejos en fragmentos uniformes, sino facilitar decisiones sobre qué información merece atención adicional. Una síntesis que conserva contexto, permite regresar a la fuente y distingue resultados de interpretaciones puede reducir el esfuerzo inicial sin empobrecer el análisis. Cuando esas condiciones no se cumplen, la comodidad del resumen puede terminar desplazando precisamente los matices que hacen científicamente significativa una publicación.

3.4 Comparación simultánea de múltiples investigaciones

La comparación simultánea de múltiples investigaciones permite examinar un conjunto de estudios mediante criterios comunes y detectar convergencias, diferencias y patrones que serían difíciles de reconocer en una lectura aislada. Esta tarea adquiere especial importancia cuando el corpus incluye numerosas publicaciones con diseños, poblaciones y resultados distintos. Tang et al. (2023) muestran que los modelos de lenguaje pueden apoyar la síntesis de evidencia científica, lo que abre la posibilidad de organizar información procedente de varios documentos. Sin embargo, comparar requiere preservar el contexto metodológico de cada estudio.

La utilidad de la comparación depende de definir previamente qué dimensiones serán contrastadas. Objetivos, diseño metodológico, características de la muestra, variables, procedimientos y resultados pueden constituir criterios de análisis, pero no todos tienen la misma relevancia en cada revisión. Ofori et al. (2024) muestran que la automatización de revisiones sistemáticas puede apoyar tareas de extracción y clasificación documental. Cuando esas operaciones se estructuran de forma consistente, es posible construir matrices comparables que permitan reconocer diferencias sustantivas sin reducir cada investigación a un conjunto descontextualizado de datos.

La heterogeneidad constituye uno de los principales problemas de este tipo de análisis. Dos estudios pueden abordar un fenómeno semejante y, al mismo tiempo, utilizar definiciones,

instrumentos o diseños incompatibles entre sí. Torre et al. (2023) muestran que la automatización de revisiones requiere organizar distintas etapas del procesamiento documental de forma coherente. Aplicado a la comparación, esto implica evitar equivalencias forzadas entre categorías que solo parecen similares. La inteligencia artificial puede detectar coincidencias textuales, pero la comparabilidad científica depende de que exista una relación metodológica o conceptual justificable.

Otro desafío aparece cuando los resultados se expresan mediante formatos distintos. Algunas publicaciones presentan efectos cuantitativos, otras describen asociaciones y otras desarrollan interpretaciones cualitativas. Doneva et al. (2024) examinan el uso de grandes modelos de lenguaje para procesar, analizar y sintetizar textos biomédicos, mostrando que estas tecnologías pueden integrar información heterogénea. No obstante, esa capacidad de integración debe utilizarse con cautela. Convertir resultados distintos en una formulación homogénea puede ocultar diferencias importantes en la naturaleza y alcance de la evidencia disponible.

La comparación también puede orientarse a identificar patrones recurrentes entre investigaciones. Gilbert et al. (2024) muestran que los modelos aumentados pueden funcionar como curadores de información cuando mantienen vínculos con fuentes verificables. Esta propiedad resulta útil para comparar varios estudios, porque permite asociar cada afirmación con el documento que la respalda. El análisis simultáneo gana rigor cuando conserva esa trazabilidad, especialmente en conjuntos extensos donde una síntesis general puede dificultar la identificación del origen específico de cada resultado o diferencia observada.

Tabla 8
Comparación de dimensiones en el análisis simultáneo de múltiples investigaciones

Dimensión	Elemento que se compara	Aporte analítico	Riesgo de comparación
Objetivos	Preguntas o propósitos de los estudios	Identificar coincidencias y diferencias de enfoque	Comparar trabajos que responden a problemas distintos
Metodología	Diseño, técnicas y procedimientos	Reconocer patrones y variaciones metodológicas	Tratar diseños diferentes como equivalentes
Muestra o corpus	Características de las unidades analizadas	Valorar alcance y comparabilidad	Ignorar diferencias contextuales relevantes
Resultados	Hallazgos reportados	Detectar convergencias o discrepancias	Separar resultados de sus condiciones metodológicas

Dimensión	Elemento que se compara	Aporte analítico	Riesgo de comparación
Limitaciones	Restricciones reconocidas por cada estudio	Ajustar el peso interpretativo de la evidencia	Omitir factores que condicionan las conclusiones

Nota. La tabla organiza dimensiones que permiten comparar investigaciones sin perder de vista las condiciones que afectan la interpretación de sus resultados.

La comparación estructurada facilita reconocer cuándo una diferencia entre estudios es sustantiva y cuándo puede explicarse por decisiones metodológicas. Tang et al. (2023) evidencian que la síntesis de evidencia exige conservar información suficiente para interpretar adecuadamente los resultados. Dos publicaciones pueden ofrecer conclusiones distintas porque trabajan con poblaciones, instrumentos o contextos diferentes. Si esas condiciones se eliminan durante la automatización, la discrepancia puede presentarse como contradicción científica cuando en realidad responde a diseños que no son directamente comparables.

La inteligencia artificial puede contribuir a detectar relaciones entre resultados dispersos, pero su capacidad de agrupar semejanzas debe estar acompañada por mecanismos de revisión. Ofori et al. (2024) muestran que las técnicas automatizadas pueden reducir la carga de trabajo en revisiones extensas, aunque la interpretación sigue dependiendo del diseño del proceso. En la comparación simultánea, esto supone revisar casos donde el sistema agrupa estudios por similitud textual pese a que sus preguntas o métodos difieren de forma significativa. La proximidad semántica no siempre implica equivalencia analítica.

La comparación de múltiples investigaciones también puede revelar vacíos en el conjunto documental. Una matriz puede mostrar que determinados contextos, poblaciones o métodos aparecen escasamente representados, mientras otros concentran gran parte de la evidencia. Torre et al. (2023) permiten entender este resultado como parte de un flujo estructurado de revisión, donde la organización de los documentos favorece una lectura más amplia del corpus. El interés no radica solo en señalar coincidencias, sino en identificar cómo se distribuye la evidencia y dónde persisten zonas con información insuficiente.

El valor de la comparación simultánea reside en hacer visibles diferencias que una síntesis excesivamente general podría borrar. Un sistema automatizado puede ordenar decenas de estudios en poco tiempo, pero la calidad del análisis depende de conservar aquello que distingue a cada investigación. La comparación se vuelve metodológicamente útil cuando permite contrastar sin homogeneizar, relacionar sin confundir y sintetizar sin perder las

condiciones que dan sentido a los resultados. Esa disciplina interpretativa evita que la eficiencia computacional sustituya la lectura crítica del corpus.

3.5 Clasificación temática y categorización automatizada

La clasificación temática automatizada permite organizar grandes colecciones de artículos según patrones de contenido que pueden identificarse mediante técnicas de procesamiento del lenguaje natural y aprendizaje automático. Su utilidad aparece cuando la revisión manual resulta insuficiente para manejar corpus extensos y heterogéneos. Chagnon et al. (2024) muestran que los modelos de tópicos aplicados a artículos científicos pueden detectar estructuras temáticas recurrentes dentro de un conjunto documental. Esta capacidad facilita ordenar la literatura y reconocer agrupamientos que no siempre son visibles mediante categorías definidas previamente por el investigador.

La categorización puede seguir dos lógicas distintas. Una parte de categorías establecidas con anterioridad y asigna cada documento a una clase conocida; otra intenta descubrir agrupamientos emergentes a partir de relaciones presentes en los textos. Wu et al. (2024) muestran que los modelos neuronales de tópicos permiten identificar patrones latentes sin depender exclusivamente de etiquetas predefinidas. La diferencia es metodológicamente importante, porque clasificar según categorías conocidas responde a un marco previo, mientras descubrir temas implica interpretar estructuras que el sistema identifica a partir del propio corpus.

La calidad de una clasificación depende de cómo se representa el contenido de los documentos. Un sistema puede utilizar palabras, expresiones, vectores semánticos o combinaciones de diferentes rasgos para estimar proximidad temática. Chagnon et al. (2024) examinan el comportamiento de modelos de tópicos en artículos científicos y muestran la necesidad de evaluar su desempeño en corpus especializados. La existencia de vocabulario técnico, términos polisémicos y expresiones propias de cada disciplina puede modificar la forma en que los documentos son agrupados y, con ello, la interpretación posterior del campo.

La categorización temática también puede emplearse durante el cribado de literatura. Dennstädt et al. (2024) analizan el uso de grandes modelos de lenguaje para clasificar títulos y resúmenes en revisiones, mostrando su potencial para apoyar decisiones iniciales sobre inclusión y exclusión. Este tipo de aplicación no busca necesariamente construir un mapa completo del campo, sino separar documentos según su posible relevancia. La tarea parece sencilla, pero una

categoría mal definida puede producir exclusiones sistemáticas de trabajos pertinentes y alterar el corpus que será analizado posteriormente.

Cuando el objetivo consiste en descubrir temas, la interpretación de las categorías generadas exige especial cuidado. Wu et al. (2024) señalan que los modelos de tópicos presentan desafíos de evaluación e interpretabilidad, porque los grupos identificados estadísticamente no siempre coinciden con categorías conceptuales claras. Un conjunto de términos frecuentes puede sugerir un tema, pero su significado depende de cómo aparecen relacionados dentro de los documentos. La etiqueta final suele ser asignada por el investigador, de modo que el componente interpretativo permanece incluso cuando el agrupamiento se produce automáticamente.

La validación permite determinar si las categorías obtenidas describen de manera razonable el corpus. Vázquez et al. (2022) muestran que los modelos de tópicos científicos pueden evaluarse mediante análisis de grafos y metadatos documentales, combinando distintas fuentes de información para examinar su coherencia. Este procedimiento evita aceptar un agrupamiento únicamente porque el algoritmo produce una separación clara entre grupos. Una categoría adquiere mayor solidez cuando presenta consistencia textual, relaciones estructurales identificables y correspondencia con características relevantes de los documentos que la integran.

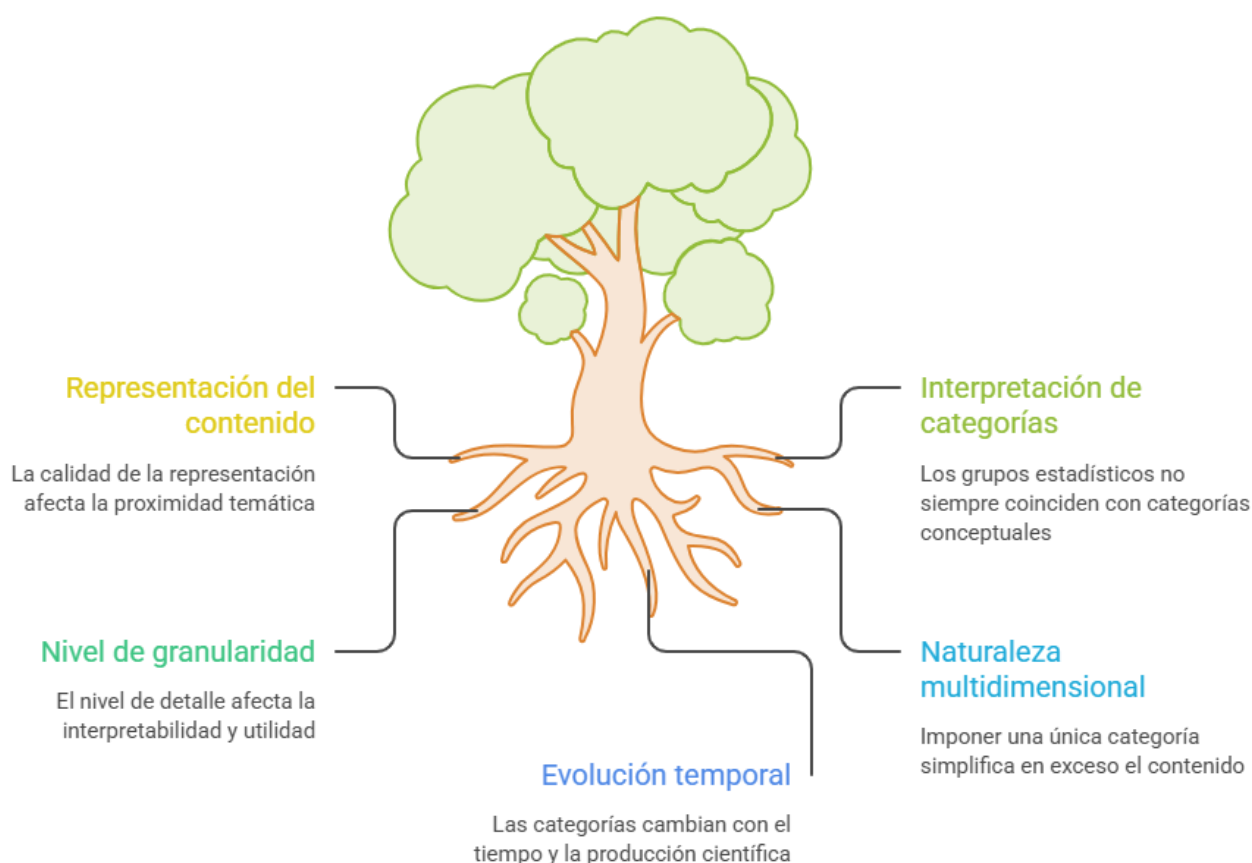
La clasificación automatizada puede operar en distintos niveles de granularidad. Un corpus puede organizarse en áreas amplias o dividirse en subtemas mucho más específicos, dependiendo del propósito analítico. Doneva et al. (2024) examinan aplicaciones de grandes modelos de lenguaje para procesar y analizar textos biomédicos, mostrando la diversidad de tareas que pueden realizarse sobre información científica. Elegir un nivel excesivamente general puede ocultar diferencias importantes, mientras una fragmentación demasiado detallada produce categorías difíciles de interpretar y reduce la utilidad del esquema para comprender la estructura documental.

Otro problema aparece cuando un artículo pertenece razonablemente a más de un tema. Las investigaciones científicas suelen combinar conceptos, métodos y problemas que atraviesan fronteras disciplinares, por lo que imponer una única categoría puede simplificar en exceso su contenido. Chagnon et al. (2024) permiten comprender que los modelos temáticos pueden representar distribuciones de tópicos dentro de los documentos y no únicamente asignaciones rígidas. Esta característica resulta útil para conservar la naturaleza multidimensional de ciertas

publicaciones y reconocer zonas de contacto entre áreas que comparten parte de su vocabulario o problemática.

La evolución temporal del corpus también puede modificar las categorías. Un término que inicialmente aparece asociado con un tema puede desplazarse hacia nuevas aplicaciones conforme cambia la producción científica. Vázquez et al. (2022) muestran que los metadatos pueden complementar la interpretación de estructuras temáticas, ofreciendo información adicional para comprender cómo se distribuyen los documentos. Incorporar la dimensión temporal permite observar si una categoría permanece estable, se fragmenta o comienza a relacionarse con otros temas. La clasificación deja así de ser estática y puede convertirse en una herramienta para estudiar transformación científica.

Figura 6
Clasificación temática automatizada ineficaz



Nota. La figura identifica factores que pueden reducir la precisión y utilidad de la clasificación temática automatizada de la literatura científica.

La automatización puede facilitar la comparación entre grandes corpus, pero las categorías solo son comparables cuando mantienen criterios suficientemente consistentes. Wu et al. (2024) advierten que los modelos de tópicos dependen de decisiones relacionadas con representación,

arquitectura y evaluación. Si dos colecciones se analizan con procedimientos distintos, las diferencias observadas pueden provenir del método y no necesariamente del contenido científico. Para realizar comparaciones válidas conviene documentar cómo se generaron las categorías y qué parámetros determinaron el nivel de agrupamiento obtenido en cada análisis.

El uso de categorías automatizadas también puede introducir una falsa sensación de exhaustividad. Un sistema siempre puede asignar documentos a grupos, incluso cuando la estructura temática del corpus es ambigua. Dennstädt et al. (2024) muestran que los modelos de lenguaje pueden apoyar tareas de clasificación, pero su desempeño debe evaluarse según el objetivo específico de la revisión. La existencia de una etiqueta no garantiza que el contenido haya sido comprendido correctamente. Los documentos fronterizos, multidisciplinarios o terminológicamente atípicos requieren una revisión más cuidadosa antes de aceptar su ubicación definitiva.

La clasificación temática resulta más útil cuando organiza la complejidad sin eliminarla. Las categorías deben facilitar la exploración del corpus, permitir localizar conjuntos relacionados y orientar preguntas posteriores, pero no convertir divisiones provisionales en fronteras rígidas del conocimiento. Un esquema automatizado puede mostrar dónde se concentran ciertos temas y qué documentos participan de varios grupos, siempre que conserve espacio para revisar casos ambiguos. La utilidad científica aparece cuando la categorización funciona como instrumento de análisis y no como sustituto de la interpretación conceptual de los textos.

3.6 Síntesis de evidencia mediante inteligencia artificial

La síntesis de evidencia mediante inteligencia artificial busca integrar información procedente de múltiples estudios en una representación coherente que permita reconocer hallazgos coincidentes, diferencias y relaciones relevantes. No se limita a resumir documentos por separado, sino que exige combinar resultados sin perder las condiciones metodológicas que explican cada hallazgo. Tang et al. (2023) muestran que los modelos de lenguaje pueden apoyar la síntesis de evidencia médica, pero su utilidad depende de conservar fidelidad respecto de las fuentes y de evitar formulaciones que excedan el contenido realmente reportado.

La automatización puede reducir parte del trabajo asociado con la revisión de grandes conjuntos documentales, especialmente cuando la información debe ser localizada, estructurada y comparada antes de construir una síntesis. Ofori et al. (2024) describen aplicaciones del procesamiento del lenguaje natural, aprendizaje automático y aprendizaje profundo en revisiones sistemáticas. Estas técnicas pueden facilitar la organización de

resultados, pero la síntesis exige algo más que reunir fragmentos semejantes. Es necesario determinar si los estudios comparados responden a preguntas compatibles y si sus resultados pueden interpretarse dentro de un mismo marco analítico.

La integración automática enfrenta dificultades cuando los documentos presentan heterogeneidad metodológica. Dos investigaciones pueden estudiar fenómenos cercanos y obtener resultados diferentes debido a variaciones en diseño, población, instrumentos o condiciones de observación. Torre et al. (2023) muestran que la automatización de revisiones científicas requiere articular varias etapas de tratamiento documental. En la síntesis, esta articulación obliga a conservar información sobre cómo se produjo cada resultado. Eliminar esas diferencias puede generar una conclusión artificialmente uniforme y ocultar factores necesarios para comprender por qué la evidencia no coincide completamente.

Los modelos de lenguaje también pueden contribuir a identificar regularidades dentro de conjuntos amplios de publicaciones. Krenn et al. (2022) plantean que la inteligencia artificial puede apoyar procesos vinculados con la comprensión científica, aunque esta posibilidad requiere distinguir entre reconocimiento de patrones e interpretación del conocimiento. Una coincidencia frecuente entre estudios puede resultar relevante, pero no basta para establecer una conclusión sólida. La síntesis necesita valorar la naturaleza de la evidencia, las condiciones en que fue obtenida y la relación entre resultados que aparentemente apuntan en una misma dirección.

La trazabilidad constituye una condición especialmente importante cuando la síntesis es generada o apoyada por sistemas automáticos. Gilbert et al. (2024) muestran la relevancia de vincular modelos de lenguaje con información recuperada de fuentes verificables para reducir respuestas no sustentadas. En una síntesis científica, cada afirmación integrada debería poder relacionarse con los documentos que la respaldan. Esta conexión permite revisar discrepancias, comprobar interpretaciones y evitar que una formulación generada por el sistema adquiera autonomía respecto de las evidencias originales que dieron lugar a su construcción.

Tabla 9
Dimensiones para evaluar una síntesis de evidencia asistida por inteligencia artificial

Dimensión	Pregunta analítica	Función de la IA	Riesgo principal
Correspondencia entre estudios	¿Los trabajos responden a problemas comparables?	Detectar similitudes temáticas y metodológicas	Integrar investigaciones que no son equivalentes

Dimensión	Pregunta analítica	Función de la IA	Riesgo principal
Conservación del contexto	¿Se mantienen las condiciones de cada resultado?	Extraer variables y atributos relevantes	Generalizar hallazgos fuera de su alcance
Identificación de convergencias	¿Qué resultados coinciden entre estudios?	Reconocer patrones recurrentes	Confundir frecuencia con solidez de evidencia
Tratamiento de discrepancias	¿Qué resultados difieren y por qué?	Localizar resultados divergentes	Suavizar contradicciones durante la síntesis
Trazabilidad	¿Puede verificarse el origen de cada afirmación?	Vincular síntesis y documentos fuente	Generar afirmaciones sin respaldo identificable

Nota. La tabla organiza criterios para diferenciar una síntesis documental verificable de una integración automática que pierde contexto o procedencia.

La identificación de convergencias debe ir acompañada del examen de las diferencias. Tang et al. (2023) muestran que la síntesis automatizada puede ser evaluada según su capacidad para representar correctamente la evidencia disponible. Si un modelo privilegia los resultados más frecuentes, puede relegar hallazgos minoritarios que resultan importantes para explicar heterogeneidad o incertidumbre. La síntesis científica no debe buscar una unanimidad artificial, sino mostrar qué resultados presentan apoyo recurrente, cuáles se apartan del patrón y qué condiciones podrían estar asociadas con esas diferencias.

La automatización también puede facilitar la extracción de información comparable antes de redactar una síntesis. Ofori et al. (2024) muestran que distintas tareas de una revisión pueden recibir apoyo mediante técnicas de inteligencia artificial, lo que permite estructurar objetivos, métodos y resultados en formatos más manejables. Esta organización ayuda a detectar relaciones entre estudios, pero la interpretación final requiere revisar si las categorías utilizadas respetan el sentido original de los documentos. Una matriz ordenada puede ser técnicamente consistente y, sin embargo, representar de forma insuficiente conceptos complejos o resultados condicionados.

Mata et al. (2024) sitúan la inteligencia artificial dentro de diferentes actividades de la investigación científica, lo que permite comprender la síntesis como parte de un proceso más amplio de apoyo tecnológico. Su incorporación puede mejorar la capacidad de procesar literatura extensa, siempre que no se convierta en una sustitución del juicio crítico. La decisión sobre qué evidencia tiene mayor peso, qué diferencias son metodológicamente relevantes y qué afirmaciones pueden sostenerse a partir del conjunto documental requiere criterios que no deberían depender exclusivamente de una salida automatizada.

La síntesis asistida por inteligencia artificial puede ser especialmente útil cuando permite pasar de un corpus extenso a una estructura argumentativa verificable. Su valor disminuye cuando elimina incertidumbres, combina resultados incompatibles o presenta inferencias como si fueran hallazgos directos de las fuentes. Una buena síntesis debe conservar diferencias suficientes para que el lector comprenda la diversidad de la evidencia disponible. La automatización contribuye cuando ayuda a organizar esa complejidad, no cuando la reduce hasta producir una conclusión más clara que los estudios que pretende representar.

3.7 Riesgos y límites de la lectura automatizada de literatura científica

La lectura automatizada de literatura científica amplía la capacidad de procesar grandes colecciones documentales, pero también introduce riesgos que afectan la interpretación y la confiabilidad de los resultados. Krenn et al. (2022) señalan que la inteligencia artificial puede contribuir a tareas vinculadas con la comprensión científica, aunque esta posibilidad no equivale a una comprensión plena del significado de los textos. Los modelos pueden reconocer patrones, resumir contenidos y establecer relaciones, pero siguen dependiendo de representaciones computacionales que no capturan necesariamente toda la complejidad conceptual de una investigación.

Uno de los límites más visibles corresponde a las alucinaciones, entendidas como la generación de información aparentemente plausible que no está respaldada por las fuentes procesadas. Gao et al. (2023) muestran que los sistemas generativos pueden producir resúmenes científicos con una apariencia lingüística difícil de distinguir de textos reales. Esta fluidez puede inducir a aceptar afirmaciones incorrectas cuando no se contrasta la salida con el documento original. La calidad formal del lenguaje, por tanto, no constituye un indicador suficiente de fidelidad científica ni de correspondencia con la evidencia disponible.

El riesgo aumenta cuando el sistema incorpora detalles inexistentes o modifica el sentido de resultados presentes en los artículos. Chelli et al. (2024) documentan problemas de alucinación y precisión de referencias en herramientas generativas utilizadas para tareas relacionadas con revisiones sistemáticas. Aunque la lectura automatizada puede acelerar la exploración inicial, sus resultados requieren comprobación antes de ser utilizados en una síntesis. Una interpretación errónea introducida en las primeras etapas puede propagarse posteriormente hacia clasificaciones, comparaciones o conclusiones construidas sobre un corpus de gran tamaño.

La condensación excesiva constituye otra limitación. Un modelo puede reducir un artículo a sus objetivos y hallazgos principales, pero omitir condiciones metodológicas necesarias para comprender el alcance de esos resultados. Tang et al. (2023), al evaluar modelos de lenguaje en la síntesis de evidencia médica, muestran que la calidad de la salida debe valorarse por su fidelidad y utilidad informativa. Cuando desaparecen restricciones, características de la muestra o diferencias entre diseños, el resumen puede parecer más concluyente que el estudio original y favorecer interpretaciones científicamente injustificadas.

La lectura automatizada también puede perder matices argumentativos que dependen de relaciones entre distintas secciones del documento. Una limitación metodológica puede modificar la interpretación de un resultado, mientras una discusión puede introducir reservas que no aparecen en el resumen. Krenn et al. (2022) permiten distinguir entre procesar grandes cantidades de información y alcanzar una comprensión científica suficientemente profunda. El problema no reside únicamente en lo que el sistema identifica, sino en aquello que puede omitir cuando transforma una argumentación compleja en una representación más breve y estructurada.

La explicabilidad constituye un límite adicional cuando no resulta posible comprender por qué un sistema seleccionó determinadas frases, documentos o relaciones. Zárate (2022) examina la explicabilidad como un componente necesario para interpretar decisiones producidas por inteligencia artificial. Aplicada a la lectura científica, esta exigencia implica poder reconocer qué información sustentó una clasificación o síntesis. Si el procedimiento permanece opaco, resulta difícil evaluar si la salida refleja el contenido del corpus o si responde a patrones internos del modelo que no guardan suficiente relación con la pregunta investigada.

La supervisión humana adquiere especial importancia cuando la herramienta interviene en decisiones que afectan qué literatura será leída o descartada. Castillo (2023) subraya la necesidad de mantener control humano cuando ChatGPT se utiliza como apoyo en la escritura científica. El mismo principio puede aplicarse a la lectura automatizada: delegar completamente el cribado o la interpretación puede producir exclusiones difíciles de detectar. La revisión humana permite recuperar documentos fronterizos, comprobar resultados inesperados y valorar matices que un sistema puede clasificar de manera insuficiente por falta de contexto disciplinar.

Los problemas de reproducibilidad también afectan el uso de modelos para procesar literatura. Kapoor y Narayanan (2023) muestran que la investigación basada en aprendizaje automático

puede enfrentar dificultades cuando existen filtraciones de información o procedimientos insuficientemente controlados. En lectura automatizada, los resultados pueden cambiar según la versión del modelo, el corpus disponible o la forma de formular las instrucciones. Si estas condiciones no se documentan, resulta difícil reconstruir cómo se obtuvo una clasificación, una síntesis o una comparación específica, reduciendo la posibilidad de auditar el proceso.

La interpretación de grandes colecciones puede verse igualmente condicionada por sesgos presentes en los datos y en los sistemas utilizados. Penabad et al. (2024) destacan la necesidad de transparencia al reportar el uso de inteligencia artificial en contextos científico-académicos. Esta transparencia resulta necesaria porque una herramienta puede favorecer ciertos tipos de documentos, vocabularios o patrones de publicación sin que el usuario lo perciba directamente. Una lectura automatizada aparentemente exhaustiva puede, por tanto, reproducir desigualdades de representación presentes en el corpus o introducir nuevas formas de selección algorítmica.

Otro riesgo aparece cuando la eficiencia tecnológica modifica los hábitos de lectura del investigador. La disponibilidad de resúmenes, clasificaciones y respuestas inmediatas puede favorecer una dependencia excesiva de representaciones secundarias del contenido. Gilbert et al. (2024) muestran que los modelos orientados a la curación de información requieren mecanismos que vinculen las respuestas con fuentes verificables. Mantener ese vínculo es decisivo para evitar que la síntesis generada se convierta en la única vía de acceso al documento. La fuente original debe permanecer disponible para contrastar afirmaciones, métodos y límites.

El principal límite de la lectura automatizada aparece cuando se intenta utilizarla para sustituir decisiones interpretativas que dependen del contexto científico. Un sistema puede señalar patrones recurrentes, pero no determina automáticamente qué diferencias son metodológicamente relevantes ni qué resultados merecen mayor peso. La utilidad de estas herramientas aumenta cuando se reservan para tareas de organización, exploración y apoyo, mientras la interpretación permanece sometida a revisión. El problema no es procesar más literatura, sino hacerlo sin perder las condiciones que permiten comprender por qué una evidencia significa lo que significa.

Capítulo

4.

Detección de tendencias, vacíos y contradicciones científicas

Jose Fernando Toala Mosquera



4.1 Identificación de tendencias en la producción científica

La identificación de tendencias en la producción científica busca reconocer cambios sostenidos en los temas, enfoques y relaciones que estructuran un campo de conocimiento. Cuando el volumen de publicaciones crece de forma constante, la lectura individual resulta insuficiente para observar patrones generales. Wu et al. (2024) muestran que los modelos neuronales de tópicos permiten identificar estructuras temáticas latentes en grandes colecciones textuales. Su aplicación a literatura científica facilita detectar áreas con presencia creciente, transformaciones en el vocabulario y desplazamientos temáticos que pueden pasar inadvertidos en revisiones documentales convencionales.

Una tendencia no debe confundirse con la simple frecuencia de aparición de un término. El aumento de determinadas palabras puede responder a cambios terminológicos sin que exista una transformación sustantiva del campo. Chagnon et al. (2024) muestran que los modelos de tópicos pueden utilizarse para reconocer configuraciones temáticas en artículos científicos, permitiendo trabajar con relaciones entre conjuntos de términos y documentos. Esta aproximación ofrece una lectura más rica que el conteo aislado, porque permite observar cómo distintas expresiones se organizan alrededor de temas que adquieren mayor o menor presencia dentro del corpus.

La dimensión temporal resulta fundamental para distinguir una tendencia de una concentración circunstancial de publicaciones. Vázquez et al. (2022) muestran que los modelos temáticos pueden complementarse con metadatos del corpus, lo que permite examinar cómo determinadas estructuras se distribuyen según diferentes atributos documentales. Cuando el año de publicación forma parte del análisis, es posible observar si un tema mantiene continuidad, crece progresivamente o pierde presencia. Esta lectura evita interpretar como tendencia un aumento breve producido por coyunturas específicas o por cambios limitados en la cobertura documental.

Los mapas bibliométricos ofrecen otra vía para examinar la evolución de la producción científica. Haunschild y Bornmann (2024) muestran que los datos de OpenAlex pueden emplearse para construir mapas globales de la ciencia en distintos niveles de análisis. Estas representaciones permiten situar conjuntos de publicaciones dentro de estructuras más amplias y observar cambios en su distribución relativa. La utilidad de este procedimiento aumenta cuando el análisis combina volumen de producción con relaciones entre áreas, pues un tema puede crecer no solo en cantidad, sino también en conexiones con otros campos.

La interpretación de tendencias exige atender a la escala utilizada. Un crecimiento visible dentro de una especialidad puede representar una variación menor cuando se compara con la producción total del área. Ávila (2024) muestra que los grafos de conocimiento permiten organizar y visualizar datos de investigación mediante relaciones estructuradas. Esta capacidad facilita cambiar el nivel de observación y examinar cómo un tema se sitúa respecto de otros. Sin esa referencia, existe el riesgo de considerar emergente aquello que únicamente ocupa una posición destacada dentro de un subconjunto documental reducido.

Tabla 10
Técnicas y criterios para identificar tendencias en la producción científica

Dimensión	Evidencia observable	Aporte analítico	Riesgo de interpretación
Frecuencia temática	Incremento de términos o tópicos	Detectar concentración creciente de atención	Confundir popularidad terminológica con cambio científico
Evolución temporal	Variación de temas por periodos	Reconocer crecimiento, estabilidad o descenso	Interpretar fluctuaciones breves como tendencias consolidadas
Relaciones entre áreas	Nuevas conexiones entre clústeres o conceptos	Identificar convergencias temáticas	Atribuir integración científica a vínculos débiles
Posición en mapas	Cambios en centralidad o proximidad	Observar reconfiguración del campo	Equiparar centralidad con relevancia científica
Persistencia	Continuidad de un patrón en varios periodos	Diferenciar tendencias de fenómenos coyunturales	Ignorar cambios recientes aún no estabilizados

Nota. La tabla organiza criterios que permiten distinguir una tendencia científica de variaciones puntuales o efectos derivados de la representación utilizada.

La combinación de dimensiones reduce la posibilidad de interpretar cambios aparentes como transformaciones reales del campo. Wu et al. (2024) muestran que los modelos de tópicos permiten estudiar estructuras latentes, pero su lectura requiere considerar cómo evolucionan y qué consistencia mantienen. Un tema que aparece con alta frecuencia en un único periodo puede responder a una concentración temporal, mientras otro con crecimiento gradual puede mostrar un cambio más sostenido. La persistencia ofrece, por tanto, una condición adicional para valorar la estabilidad de un patrón observado.

Las tendencias también pueden manifestarse en la aparición de conexiones entre áreas antes relativamente separadas. Haunschild y Bornmann (2024) permiten examinar estas relaciones mediante mapas que sitúan unidades científicas dentro de estructuras globales. Cuando un

conjunto de publicaciones comienza a ocupar posiciones intermedias entre campos, puede indicar una convergencia temática o metodológica. Sin embargo, esa interpretación requiere revisar el contenido de los documentos, porque una proximidad bibliométrica puede derivarse de citas compartidas sin que exista todavía una integración conceptual suficientemente desarrollada.

La detección automatizada puede facilitar la identificación de patrones, pero la selección del corpus condiciona fuertemente lo que se observa. Vázquez et al. (2022) muestran la importancia de utilizar metadatos y estructuras del corpus para validar modelos temáticos. Si una base de datos cambia su cobertura, incorpora nuevas revistas o presenta diferencias disciplinares, el crecimiento aparente de un tema puede deberse parcialmente a esas modificaciones. El análisis de tendencias debe separar, en la medida de lo posible, cambios en la producción científica de cambios en la fuente documental utilizada.

Una tendencia adquiere valor analítico cuando puede sostenerse mediante una secuencia temporal, una estructura temática coherente y evidencia documental suficiente. El objetivo no consiste en señalar qué tema aparece más veces, sino en comprender cómo cambia su posición dentro de un campo y qué relaciones nuevas establece con otras líneas de investigación. Esa lectura exige combinar automatización con contraste contextual. Los sistemas pueden hacer visible el patrón, pero corresponde al análisis científico determinar si representa una transformación persistente o solo una variación circunstancial del corpus.

4.2 Evolución temporal de temas y líneas de investigación

El análisis temporal de temas científicos permite observar cómo cambian las prioridades de investigación, qué conceptos ganan presencia y cuáles pierden centralidad dentro de un campo. No se limita a contar publicaciones por año, sino que busca identificar transformaciones en la estructura temática de un corpus. Vázquez et al. (2022) muestran que los modelos de tópicos pueden analizarse junto con metadatos documentales, lo que ofrece una vía para relacionar patrones temáticos con periodos de publicación y examinar su comportamiento a través del tiempo.

Una línea de investigación puede mantenerse estable durante varios años y, aun así, modificar su vocabulario, sus problemas centrales o sus conexiones con otras áreas. Wu et al. (2024) señalan que los modelos neuronales de tópicos permiten representar estructuras temáticas latentes en colecciones textuales, lo que facilita estudiar variaciones que no siempre son visibles mediante categorías fijas. Esta capacidad resulta especialmente útil cuando un campo

evoluciona sin abandonar completamente sus temas previos, sino reorganizando conceptos existentes y generando nuevas asociaciones entre ellos.

El tiempo también permite distinguir entre temas consolidados, emergentes y transitorios. Un aumento breve en el número de publicaciones puede responder a una coyuntura específica, mientras un crecimiento sostenido durante varios periodos sugiere una incorporación más estable dentro de la agenda científica. Haunschild y Bornmann (2024) muestran que los mapas bibliométricos pueden representar la ciencia a diferentes escalas, lo que permite observar desplazamientos relativos entre áreas. Esta perspectiva ayuda a valorar si un tema realmente amplía su presencia o simplemente crece al mismo ritmo que el resto del campo.

Los cambios temáticos pueden producirse mediante expansión, fragmentación o convergencia. Una línea consolidada puede diversificarse en subtemas más específicos, mientras dos áreas antes separadas pueden comenzar a compartir conceptos y documentos. Chagnon et al. (2024) muestran que los modelos de tópicos aplicados a artículos científicos permiten identificar agrupamientos temáticos a partir de patrones de contenido. Cuando estos agrupamientos se analizan por periodos, es posible observar cómo se reorganiza la estructura interna del corpus y qué relaciones comienzan a adquirir mayor consistencia.

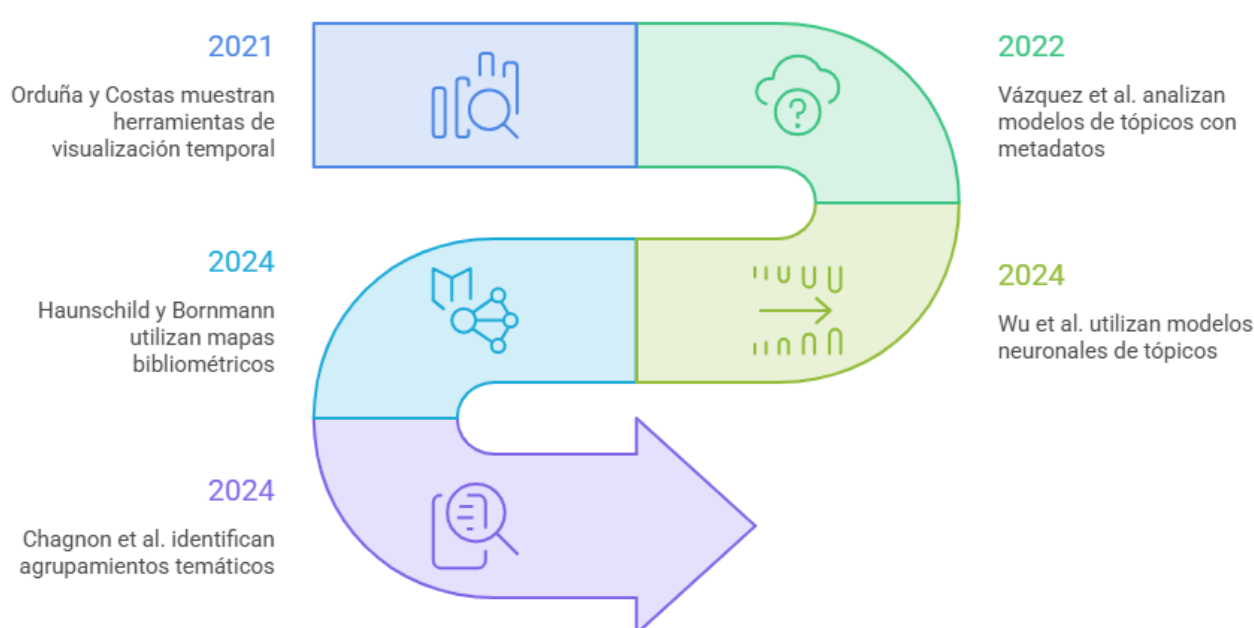
La visualización temporal aporta una lectura complementaria porque permite representar desplazamientos que serían difíciles de reconocer en tablas extensas. Orduña y Costas (2021) muestran la presencia de herramientas como VOSviewer en prácticas de análisis científico basadas en relaciones. Este tipo de software puede facilitar la comparación de estructuras en diferentes periodos y mostrar cambios en agrupamientos, conexiones o posiciones relativas. Sin embargo, la interpretación debe considerar que las diferencias visuales pueden depender también de parámetros técnicos y no únicamente de transformaciones reales en la producción científica.

La evolución de una línea de investigación tampoco debe inferirse exclusivamente a partir del volumen de publicaciones. Un tema puede recibir mucha atención sin experimentar cambios conceptuales relevantes, mientras otro puede presentar pocas publicaciones, pero introducir relaciones nuevas con áreas previamente distantes. Vázquez et al. (2022) permiten ampliar el análisis mediante información estructural y metadatos, evitando reducir la evolución a una serie de frecuencias. La combinación entre contenido, relaciones y tiempo ofrece una descripción más precisa de cómo se transforma una agenda científica.

Los conceptos que funcionan como puentes entre líneas distintas merecen especial atención en el análisis temporal. Cuando una expresión comienza a aparecer en varios grupos temáticos, puede indicar un proceso de convergencia o una expansión de su campo de aplicación. Wu et al. (2024) muestran que los modelos de tópicos pueden capturar relaciones latentes entre términos y documentos, ofreciendo herramientas para detectar estas transiciones. No obstante, la aparición conjunta de conceptos debe contrastarse con los textos para comprobar si existe una relación científica sustantiva y no solo una coincidencia terminológica.

Figura 7

Evolución temporal de temas y líneas de investigación



Nota. La figura muestra la evolución de distintos enfoques utilizados para identificar, representar y analizar cambios temáticos en la producción científica a lo largo del tiempo.

Las discontinuidades también forman parte de la evolución temática. Algunas líneas disminuyen su presencia porque ciertas preguntas pierden prioridad, porque son absorbidas por enfoques más amplios o porque cambian las formas de denominarlas. Haunschild y Bornmann (2024) permiten observar estas variaciones dentro de mapas bibliométricos que sitúan unidades científicas en contextos amplios. Una reducción de visibilidad no implica necesariamente desaparición del conocimiento producido. Puede representar una reconfiguración de sus vínculos o una migración hacia nuevas categorías que requieren seguimiento más detallado.

El análisis longitudinal exige mantener comparabilidad entre periodos. Si cambian de manera sustancial las fuentes de datos, los criterios de inclusión o los procedimientos de clasificación, las diferencias observadas pueden atribuirse erróneamente a la evolución del campo. Chagnon

et al. (2024) muestran que distintos modelos pueden producir estructuras temáticas diferentes, por lo que la consistencia metodológica resulta importante al comparar intervalos temporales. Documentar el procedimiento permite distinguir entre transformaciones del corpus y variaciones introducidas por el método utilizado para representarlo.

La dimensión temporal también ayuda a identificar cuándo una línea empieza a adquirir autonomía. Un conjunto inicialmente integrado dentro de un tema amplio puede aumentar su producción, desarrollar vocabulario propio y formar conexiones internas más densas. Orduña y Costas (2021) permiten entender cómo las relaciones entre elementos científicos pueden observarse mediante estructuras de red. Al seguir estos cambios en varios periodos, es posible examinar procesos de diferenciación temática sin asumir que toda agrupación reciente constituye automáticamente una nueva área de investigación.

La utilidad del análisis temporal aparece cuando permite reconstruir procesos y no solo ordenar publicaciones cronológicamente. Observar cómo un tema se expande, se conecta con otros o cambia de posición ayuda a comprender la dinámica interna del conocimiento científico. Sin embargo, ninguna secuencia visual o estadística explica por sí misma las razones de esas transformaciones. El análisis debe volver a los documentos para interpretar qué cambios conceptuales o metodológicos sostienen la evolución observada y evitar convertir una variación cuantitativa en una narrativa histórica no demostrada.

4.3 Detección de temas emergentes mediante inteligencia artificial

La detección de temas emergentes busca identificar áreas de investigación que comienzan a ganar presencia dentro de un campo antes de alcanzar una consolidación clara. Esta tarea exige distinguir entre una novedad circunstancial y un patrón que empieza a mostrar continuidad temática, crecimiento documental y nuevas conexiones. Wu et al. (2024) muestran que los modelos neuronales de tópicos pueden revelar estructuras latentes en grandes colecciones textuales. Aplicados longitudinalmente, permiten reconocer configuraciones que incrementan su presencia y que podrían corresponder a líneas científicas en proceso de formación.

La emergencia temática no depende únicamente del aumento de publicaciones. Un tema puede adquirir relevancia porque empieza a relacionarse con conceptos procedentes de áreas distintas o porque genera agrupamientos documentales antes inexistentes. Chagnon et al. (2024) muestran que los modelos de tópicos aplicados a artículos científicos permiten descubrir patrones temáticos mediante representaciones computacionales del contenido. Esta capacidad

resulta útil para localizar agrupaciones incipientes, aunque la aparición de un nuevo clúster debe interpretarse con cautela y contrastarse con la evolución efectiva del corpus analizado.

Los metadatos aportan información necesaria para determinar si un tema emergente mantiene una trayectoria temporal reconocible. Vázquez et al. (2022) muestran que los modelos temáticos pueden validarse mediante grafos y características del corpus, lo que permite combinar contenido con información estructural. Un tópico que aparece en varios periodos, aumenta su volumen y consolida relaciones internas ofrece señales más consistentes de emergencia que otro detectado únicamente en un intervalo breve. La dimensión temporal ayuda así a separar cambios persistentes de fluctuaciones propias de la producción científica.

La inteligencia artificial también puede facilitar la detección de relaciones nuevas entre temas previamente separados. Hua et al. (2023) describen avances en algoritmos inteligentes aplicados al procesamiento científico, mostrando la ampliación de capacidades para reconocer patrones complejos. En el análisis de literatura, estas técnicas pueden señalar convergencias entre áreas mediante similitudes semánticas o relaciones documentales. Cuando una conexión comienza a repetirse y a generar producción específica, puede constituir una señal de transformación temática. Aun así, la conexión algorítmica requiere interpretación antes de considerarse una línea emergente.

La aparición de un tema nuevo puede relacionarse con cambios en problemas científicos, métodos o formas de conceptualizar fenómenos existentes. Krenn et al. (2022) plantean que la inteligencia artificial puede contribuir a la exploración y comprensión de estructuras científicas complejas. Esta capacidad puede apoyar la identificación de combinaciones conceptuales poco frecuentes o conexiones que todavía no ocupan posiciones centrales. Sin embargo, detectar novedad estadística no equivale a demostrar novedad científica, porque una combinación poco frecuente puede reflejar simplemente una variación terminológica o un corpus insuficientemente representativo.

Tabla 11
Criterios analíticos para identificar temas científicos emergentes mediante IA

Criterio	Evidencia observable	Aporte del análisis automatizado	Riesgo interpretativo
Crecimiento reciente	Aumento progresivo de publicaciones	Detectar variaciones temporales	Confundir crecimiento puntual con consolidación
Formación de clústeres	Agrupamientos nuevos o más densos	Reconocer estructuras temáticas incipientes	Sobre interpretar agrupaciones inestables

Criterio	Evidencia observable	Aporte del análisis automatizado	Riesgo interpretativo
Convergencia conceptual	Nuevas relaciones entre áreas	Identificar proximidades semánticas	Confundir coocurrencia con integración científica
Persistencia temporal	Presencia continuada en varios periodos	Diferenciar tendencias de episodios breves	Excluir temas muy recientes pero relevantes
Diversificación	Aparición de subtemas o aplicaciones	Observar expansión interna	Interpretar fragmentación como madurez temática

Nota. La tabla organiza señales que pueden utilizarse conjuntamente para valorar si un patrón detectado representa una emergencia temática consistente.

La tabla muestra que ninguna señal aislada resulta suficiente para afirmar que un tema está emergiendo. Un aumento cuantitativo puede producirse sin cambios conceptuales relevantes, mientras un clúster nuevo puede desaparecer cuando se actualiza el corpus. Vázquez et al. (2022) permiten reforzar la necesidad de contrastar resultados temáticos con información estructural y metadatos. La convergencia entre varias señales ofrece una base más sólida para interpretar el cambio, especialmente cuando el patrón mantiene continuidad y desarrolla relaciones diferenciadas respecto de áreas previamente establecidas.

Los sistemas automatizados también pueden ayudar a localizar temas aún pequeños que quedarían ocultos bajo análisis centrados en frecuencia absoluta. Wu et al. (2024) muestran que los modelos neuronales de tópicos pueden representar patrones latentes incluso cuando no dominan el corpus. Esta característica resulta útil para detectar señales tempranas, pero obliga a controlar el riesgo de sobredimensionarlas. Un grupo pequeño de documentos puede representar una innovación relevante o simplemente una variación marginal. La revisión del contenido permite establecer si existe una agenda científica reconocible detrás del agrupamiento.

La interpretación de un tema emergente requiere además observar qué tipo de producción comienza a concentrarse alrededor de él. Chagnon et al. (2024) evidencian que la estructura temática depende de cómo se representan y agrupan los documentos. Por ello, conviene examinar si el supuesto tema reúne trabajos con preguntas relacionadas, vocabulario consistente y conexiones internas suficientes. Una etiqueta generada a partir de términos frecuentes puede resultar atractiva, pero carece de valor analítico si los documentos incluidos no comparten un problema científico identificable o una trayectoria de desarrollo común.

La detección temprana resulta útil cuando orienta nuevas preguntas sin convertir señales preliminares en afirmaciones definitivas sobre el futuro de un campo. Un tema emergente

puede consolidarse, fragmentarse o perder relevancia conforme aparezcan nuevas publicaciones. La inteligencia artificial ayuda a localizar configuraciones que merecen seguimiento, pero no puede garantizar su permanencia. El análisis más prudente trata la emergencia como una condición provisional, sustentada por indicios observables y sujeta a revisión conforme cambia el corpus científico y se acumula nueva evidencia.

4.4 Identificación de vacíos de conocimiento

La identificación de vacíos de conocimiento implica reconocer preguntas, relaciones o contextos que permanecen insuficientemente estudiados dentro de un campo científico. No basta con señalar que existen pocos documentos sobre un tema; es necesario determinar qué aspecto concreto permanece sin explicación suficiente. Arias y Artigas (2022) vinculan la formulación de problemas relevantes con la identificación de brechas de investigación, lo que sitúa el vacío como resultado de un análisis crítico de la literatura y no como una ausencia inferida únicamente a partir del número de publicaciones disponibles.

Un vacío puede aparecer cuando existe abundante producción sobre un fenómeno, pero ciertos grupos, variables o condiciones reciben escasa atención. Esta distinción resulta importante porque una literatura numerosa puede generar una impresión engañosa de cobertura suficiente. Mata et al. (2024), al revisar el uso de inteligencia artificial en procesos de investigación científica, muestran la diversidad de aplicaciones que estas tecnologías pueden asumir. Un análisis automatizado de grandes corpus puede ayudar a reconocer concentraciones temáticas y, por contraste, áreas poco desarrolladas cuya limitada presencia merece una revisión más específica.

Los modelos de tópicos ofrecen una vía para explorar estas distribuciones internas. Wu et al. (2024) muestran que los modelos neuronales pueden detectar estructuras temáticas latentes dentro de colecciones documentales amplias. Cuando ciertos conceptos aparecen escasamente conectados con temas dominantes o forman agrupamientos reducidos, pueden señalar zonas potencialmente poco exploradas. Sin embargo, la baja frecuencia no constituye por sí sola evidencia de un vacío. También puede reflejar terminología diferente, especialización extrema o una delimitación inadecuada del corpus utilizado para realizar el análisis.

Los metadatos permiten precisar mejor dónde se concentra o falta la producción científica. Vázquez et al. (2022) muestran que la estructura temática puede analizarse conjuntamente con características documentales y relaciones de grafos. Esta combinación permite observar si determinados temas están asociados casi exclusivamente con ciertos periodos, contextos o tipos

de documentos. Una brecha puede surgir cuando una pregunta ha sido estudiada en una población concreta, pero permanece prácticamente ausente en otras condiciones. La ausencia adquiere significado cuando puede relacionarse con una dimensión científicamente pertinente.

También existen vacíos metodológicos. Un campo puede concentrar numerosos estudios sobre el mismo problema, pero depender de diseños, técnicas o fuentes de datos muy semejantes. Arias y Artigas (2022) permiten comprender que una brecha relevante no se limita al tema estudiado, sino que puede encontrarse en la forma en que el problema ha sido abordado. Identificar esta concentración exige comparar procedimientos y reconocer qué alternativas han recibido menor atención. La inteligencia artificial puede facilitar esa clasificación inicial, aunque la importancia metodológica del vacío requiere interpretación especializada.

Los vacíos conceptuales son más difíciles de detectar mediante frecuencia documental. Pueden aparecer cuando determinados conceptos son utilizados de forma fragmentada, cuando existen relaciones teóricas poco desarrolladas o cuando varios enfoques permanecen desconectados entre sí. Krenn et al. (2022) plantean que la inteligencia artificial puede contribuir a explorar estructuras de conocimiento y relaciones científicas complejas. Esta capacidad permite localizar conexiones potenciales entre áreas, pero no demuestra que una relación ausente constituya necesariamente una brecha relevante. Su valor depende de la posibilidad de formular un problema científicamente defendible.

El análisis temporal también ayuda a distinguir vacíos persistentes de cuestiones simplemente recientes. Una ausencia observada durante un periodo corto puede reflejar que un tema acaba de comenzar a desarrollarse, mientras una escasa atención mantenida durante varios años puede señalar una brecha más estable. Wu et al. (2024) permiten utilizar estructuras temáticas para observar cambios en grandes colecciones documentales. La dimensión temporal aporta contexto a la ausencia y evita interpretar del mismo modo un tema emergente, una línea abandonada y un problema que nunca recibió atención suficiente.

La automatización puede facilitar la localización de zonas poco representadas, pero también puede producir falsos vacíos cuando el corpus es incompleto. Vázquez et al. (2022) muestran que la validación de modelos temáticos requiere considerar la estructura y los metadatos de los documentos analizados. Si una base excluye determinadas revistas, idiomas o áreas disciplinares, la aparente ausencia de investigación puede ser consecuencia de la cobertura documental. Antes de formular una brecha, resulta necesario comprobar si la carencia pertenece realmente al conocimiento disponible o al sistema utilizado para recuperarlo.

La relevancia de un vacío tampoco se deriva automáticamente de su existencia. No toda combinación poco estudiada merece convertirse en una nueva pregunta de investigación. Arias y Artigas (2022) enfatizan la importancia de identificar brechas que conduzcan a problemas científicos relevantes. Esto exige valorar si la ausencia detectada afecta la comprensión de un fenómeno, limita la capacidad explicativa disponible o impide contrastar resultados en condiciones significativas. La inteligencia artificial puede señalar zonas poco exploradas, pero la pertinencia científica de esas zonas requiere una justificación conceptual adicional.

El análisis de grandes colecciones puede contribuir a comparar vacíos entre áreas y reconocer dónde se concentran las limitaciones del conocimiento disponible. Mata et al. (2024) muestran que la inteligencia artificial se incorpora a diversas fases de la investigación, lo que amplía las posibilidades de organizar y revisar literatura científica de manera sistemática. Cuando estas capacidades se aplican a la identificación de brechas, conviene mantener la trazabilidad entre la señal detectada y los documentos que la originan. Una ausencia debe poder comprobarse y no quedar reducida a una inferencia algorítmica.

Un vacío de conocimiento adquiere valor cuando puede formularse con suficiente precisión para orientar una investigación concreta. Expresiones generales como “faltan estudios” resultan insuficientes si no se especifica qué falta, respecto de qué evidencia y bajo qué condiciones. La automatización puede hacer visibles concentraciones, ausencias y conexiones débiles, pero la formulación final exige traducir esas señales en una pregunta científicamente justificable. El objetivo no es producir nuevas investigaciones a partir de cualquier zona vacía del mapa, sino reconocer aquellas cuya exploración puede resolver una limitación identificable del conocimiento disponible.

4.5 Reconocimiento de resultados divergentes y contradicciones científicas

La existencia de resultados divergentes forma parte del desarrollo normal del conocimiento científico y no debe interpretarse de manera inmediata como evidencia de error. Dos investigaciones pueden producir hallazgos diferentes porque emplean diseños, poblaciones, instrumentos o contextos distintos. Wadden y Lo (2021) muestran que la verificación de afirmaciones científicas exige determinar si la evidencia disponible respalda, contradice o resulta insuficiente frente a una proposición concreta. Esta lógica permite abordar la divergencia como un problema de relación entre afirmaciones y evidencias, antes que como una simple oposición entre conclusiones.

El reconocimiento automatizado de contradicciones requiere identificar primero qué afirmación sostiene cada documento. Una misma investigación puede contener resultados compatibles con ciertos estudios y discrepantes con otros, por lo que la comparación debe realizarse en unidades suficientemente precisas. Wadden et al. (2022) desarrollan un enfoque de verificación científica en dominio abierto orientado a relacionar afirmaciones con evidencia procedente de una colección amplia de resúmenes. Este tipo de procedimiento permite localizar documentos potencialmente favorables o contrarios a una proposición, aunque la clasificación obtenida necesita interpretación contextual.

La divergencia puede surgir también de diferencias metodológicas y no de una oposición sustantiva entre hallazgos. Tang et al. (2023) muestran que la síntesis de evidencia mediante modelos de lenguaje exige conservar suficiente información para representar con fidelidad los resultados de los estudios. Si dos trabajos emplean poblaciones o diseños distintos, una comparación superficial puede convertir una diferencia esperable en una supuesta contradicción. El análisis debe examinar primero las condiciones bajo las cuales se produjeron los resultados antes de determinar si existe realmente una incompatibilidad científica.

Los sistemas automatizados pueden contribuir a localizar pares o conjuntos de estudios con afirmaciones potencialmente discordantes. Wadden et al. (2022) muestran que la verificación de afirmaciones científicas puede abordarse mediante recuperación documental y clasificación de evidencia. Esta capacidad resulta útil en corpus extensos, donde revisar manualmente todas las combinaciones posibles sería poco viable. Sin embargo, detectar oposición textual no basta para establecer contradicción científica, porque dos afirmaciones pueden utilizar formulaciones distintas, referirse a poblaciones diferentes o depender de condiciones que modifican su alcance.

La contradicción también puede aparecer cuando una afirmación general es apoyada por algunos estudios y limitada por otros. Krenn et al. (2022) plantean que la inteligencia artificial puede contribuir a organizar y comprender relaciones complejas dentro del conocimiento científico. Esta posibilidad facilita examinar cómo se distribuye el apoyo a una proposición, pero no elimina la necesidad de valorar la fuerza y el contexto de cada evidencia. La frecuencia de estudios favorables no resuelve automáticamente una controversia cuando los trabajos difieren en calidad, diseño o condiciones de aplicación.

Tabla 12*Dimensiones para analizar resultados divergentes y contradicciones científicas*

Dimensión	Pregunta analítica	Aporte posible de la IA	Riesgo interpretativo
Afirmación	¿Qué sostiene exactamente cada estudio?	Extraer proposiciones comparables	Simplificar argumentos complejos
Evidencia	¿Qué resultado respalda la afirmación?	Vincular proposiciones con fragmentos documentales	Separar el resultado de su contexto
Condiciones	¿En qué población, diseño o situación se obtuvo?	Identificar atributos metodológicos	Tratar estudios distintos como equivalentes
Relación	¿La evidencia apoya, limita o contradice otra afirmación?	Clasificar vínculos entre resultados	Convertir diferencias parciales en oposición absoluta
Alcance	¿Hasta dónde puede generalizarse la contradicción?	Comparar múltiples documentos	Extrapolar más allá de lo demostrado

Nota. La tabla diferencia los niveles que deben examinarse antes de considerar que dos resultados constituyen una contradicción científica.

La clasificación entre apoyo y contradicción resulta especialmente útil cuando la afirmación está claramente delimitada. Wadden y Lo (2021) muestran que la verificación científica puede estructurarse alrededor de relaciones explícitas entre una proposición y la evidencia disponible. Sin embargo, muchas controversias reales no adoptan una forma binaria. Un estudio puede respaldar parcialmente una afirmación, limitarla a determinadas condiciones o cuestionar solo una parte de su alcance. La interpretación automatizada debe evitar transformar estos matices en etiquetas demasiado rígidas que oculten la complejidad del desacuerdo.

La síntesis automatizada puede introducir otro problema si intenta resolver divergencias produciendo una formulación intermedia aparentemente conciliadora. Tang et al. (2023) muestran que los modelos de lenguaje pueden sintetizar evidencia, pero la fidelidad de esa síntesis debe evaluarse cuidadosamente. Cuando existen resultados incompatibles, una salida que los combine sin explicitar la discrepancia puede generar una falsa impresión de consenso. La contradicción debe mantenerse visible hasta que exista una base suficiente para explicar sus causas o establecer qué condiciones diferencian los hallazgos comparados.

Los sistemas orientados a reducir alucinaciones también pueden aportar a este análisis cuando mantienen la relación entre las afirmaciones producidas y las fuentes consultadas. Gilbert et al. (2024) destacan la utilidad de modelos aumentados que funcionan como curadores de información vinculada con evidencia verificable. En el estudio de contradicciones, esta trazabilidad permite regresar a cada documento y comprobar qué se afirmó realmente. El valor

del sistema no reside en decidir cuál estudio tiene razón, sino en organizar la evidencia de forma que el desacuerdo pueda ser examinado con mayor precisión.

Una contradicción científicamente relevante debe formularse con suficiente detalle para mostrar qué resultados son incompatibles y bajo qué condiciones. Etiquetar un campo como “contradictorio” por la mera coexistencia de conclusiones diferentes simplifica en exceso el problema. La tarea analítica consiste en separar discrepancias aparentes, diferencias metodológicas y oposiciones verdaderamente sustantivas. La inteligencia artificial puede acelerar esa identificación y ordenar grandes volúmenes de evidencia, pero la contradicción solo adquiere significado cuando se reconstruye el contexto que hace que dos afirmaciones no puedan sostenerse simultáneamente.

4.6 Análisis de patrones metodológicos y conceptuales

El análisis de patrones metodológicos y conceptuales permite reconocer regularidades que atraviesan un conjunto amplio de investigaciones y que no siempre son visibles mediante una lectura aislada de cada estudio. Estas regularidades pueden aparecer en los diseños utilizados, las técnicas de recolección, las formas de definir variables o las categorías conceptuales empleadas. Chagnon et al. (2024) muestran que los modelos temáticos aplicados a literatura científica facilitan la detección de estructuras recurrentes. Su utilidad aumenta cuando esas estructuras se contrastan con decisiones metodológicas presentes en los documentos analizados.

Los patrones metodológicos pueden manifestarse mediante la reiteración de determinados diseños, poblaciones, instrumentos o procedimientos analíticos. Identificarlos permite observar qué formas de producción de evidencia predominan dentro de un campo y cuáles aparecen con menor frecuencia. Doneva et al. (2024) examinan el uso de grandes modelos de lenguaje para procesar, analizar y sintetizar textos científicos, lo que abre posibilidades para extraer información metodológica de colecciones extensas. Sin embargo, la automatización debe conservar suficiente contexto para evitar tratar como equivalentes procedimientos que solo comparten etiquetas generales.

La comparación de diseños requiere distinguir entre semejanza formal y equivalencia metodológica. Dos estudios pueden declararse cuantitativos y, sin embargo, diferir de manera importante en sus muestras, instrumentos o estrategias de análisis. De la Cruz (2017) relaciona el rigor científico con la coherencia interna del proceso investigativo, criterio que resulta útil para examinar patrones sin reducirlos a frecuencias. La repetición de un diseño no demuestra

por sí misma su adecuación; interesa observar cómo se articula con el problema estudiado y con la calidad de la evidencia producida.

Los patrones conceptuales, por su parte, se relacionan con la manera en que determinados términos, categorías o relaciones teóricas se repiten y se organizan dentro de la literatura. Acevedo et al. (2017) muestran que los modelos científicos cumplen funciones de representación y explicación, lo que permite comprender que los conceptos no operan como etiquetas aisladas. Cuando una colección documental utiliza de forma recurrente ciertas categorías para representar un fenómeno, esas regularidades pueden revelar marcos compartidos, aunque también pueden ocultar diferencias internas en la forma de definirlos y aplicarlos.

La inteligencia artificial puede apoyar esta tarea mediante clasificación temática y extracción de relaciones. Wu et al. (2024) muestran que los modelos neuronales de tópicos permiten reconocer patrones latentes en grandes colecciones textuales, mientras Vázquez et al. (2022) plantean que estas estructuras pueden contrastarse con metadatos y relaciones documentales. La combinación de ambas perspectivas resulta útil para observar si determinados conceptos aparecen asociados de manera consistente con ciertos métodos. No obstante, una asociación estadística entre tema y diseño requiere interpretación antes de considerarse una pauta científica estable.

La relación entre patrones conceptuales y metodológicos puede revelar dependencias relevantes dentro de un campo. Un determinado enfoque teórico puede aparecer vinculado de manera recurrente con ciertos diseños, poblaciones o procedimientos de análisis. Krenn et al. (2022) plantean que la inteligencia artificial puede contribuir a explorar estructuras complejas del conocimiento científico, facilitando la identificación de relaciones difíciles de detectar manualmente. El interés analítico reside en determinar si esas asociaciones responden a necesidades del problema investigado o si reflejan hábitos metodológicos consolidados dentro de una comunidad científica.

La repetición metodológica también puede señalar zonas de estabilidad, pero no debe interpretarse automáticamente como fortaleza. Un campo puede acumular numerosos estudios que utilizan procedimientos semejantes y, aun así, mantener limitaciones derivadas de esa homogeneidad. De la Cruz (2017) permite situar el rigor en la correspondencia entre decisiones metodológicas y propósito científico, más que en la mera repetición de prácticas aceptadas. Identificar una pauta dominante resulta útil precisamente porque permite preguntar qué

aspectos del fenómeno quedan menos explorados cuando la investigación recurre continuamente a estrategias similares.

Los patrones conceptuales pueden mostrar también procesos de convergencia terminológica. Cuando diferentes investigaciones comienzan a utilizar categorías semejantes, puede existir una tendencia hacia mayor uniformidad en la representación de un problema. Chagnon et al. (2024) muestran que los modelos de tópicos permiten reconocer agrupamientos temáticos en artículos científicos, ofreciendo una vía para observar esas concentraciones. Sin embargo, la convergencia del vocabulario no implica necesariamente acuerdo conceptual. Dos autores pueden utilizar el mismo término y atribuirle significados distintos, por lo que la revisión contextual sigue siendo indispensable.

El análisis automatizado facilita además reconocer combinaciones poco frecuentes entre métodos y conceptos. Doneva et al. (2024) permiten entender que los modelos de lenguaje pueden procesar grandes volúmenes de textos y apoyar tareas de comparación. Estas combinaciones atípicas pueden señalar innovaciones metodológicas, cruces disciplinares o simplemente excepciones dentro del corpus. Su interpretación exige examinar si representan una desviación significativa respecto del patrón dominante y si aportan resultados distintos. La rareza estadística, por sí sola, no basta para atribuir relevancia científica a una configuración metodológica o conceptual.

La dimensión temporal agrega profundidad al análisis de patrones. Un método puede dominar durante un periodo y luego perder presencia, mientras ciertas categorías conceptuales se transforman o se combinan con nuevas perspectivas. Vázquez et al. (2022) muestran que los metadatos documentales pueden utilizarse para contextualizar estructuras temáticas, permitiendo observar cambios a lo largo del tiempo. Esta lectura ayuda a distinguir patrones persistentes de configuraciones transitorias y evita tratar la estructura actual del campo como si hubiese permanecido estable durante toda su trayectoria de desarrollo.

El análisis de patrones también puede contribuir a identificar tensiones entre discurso conceptual y práctica metodológica. Una literatura puede declarar interés por fenómenos complejos, pero apoyarse de manera reiterada en procedimientos que capturan solo una parte de esa complejidad. Krenn et al. (2022) permiten situar este problema dentro de la relación entre inteligencia artificial y comprensión científica. La automatización puede mostrar regularidades entre conceptos y métodos, pero la interpretación crítica debe preguntar si existe

coherencia entre lo que los estudios dicen investigar y lo que sus procedimientos realmente permiten observar.

Figura 8
Equilibrando métodos y conceptos en la investigación



Nota. La figura contrasta patrones metodológicos y conceptuales como dimensiones complementarias para examinar la consistencia de una investigación.

Un patrón adquiere interés científico cuando ayuda a explicar cómo se está construyendo el conocimiento dentro de un campo y qué consecuencias tiene esa forma de construcción. La reiteración de conceptos, métodos o combinaciones entre ambos puede revelar decisiones compartidas que orientan la producción de evidencia, pero también límites que permanecen poco cuestionados. El análisis automatizado resulta útil para hacer visibles estas regularidades; su aporte más sólido aparece cuando esas señales conducen a revisar críticamente las prácticas dominantes, en lugar de convertir la frecuencia en un criterio automático de validez.

4.7 Uso de tendencias y vacíos para formular nuevas investigaciones

La formulación de nuevas investigaciones puede apoyarse en el análisis de tendencias y vacíos cuando estos se interpretan como señales de problemas todavía abiertos, transformaciones temáticas o concentraciones insuficientes de evidencia. Arias y Artigas (2022) vinculan la identificación de brechas con el planteamiento de problemas científicos relevantes, lo que impide reducir esta tarea a localizar temas poco estudiados. Una línea de investigación adquiere sentido cuando la ausencia detectada puede traducirse en una pregunta concreta, justificable y relacionada con una limitación reconocible del conocimiento disponible.

Las tendencias aportan otra clase de información, porque muestran hacia dónde se desplaza la atención científica y qué problemas comienzan a concentrar mayor producción. Sin embargo, un tema creciente no debe convertirse automáticamente en una nueva investigación. Krenn et al. (2022) plantean que la inteligencia artificial puede contribuir a explorar estructuras complejas del conocimiento científico, pero la interpretación de esas estructuras exige discernir entre novedad aparente y relevancia sustantiva. Una tendencia resulta más útil cuando permite reconocer cambios conceptuales, metodológicos o contextuales que todavía requieren explicación.

El análisis automatizado puede ayudar a relacionar ambos elementos. Un campo puede mostrar crecimiento acelerado en ciertas áreas y, al mismo tiempo, mantener vacíos en poblaciones, métodos o contextos específicos. Mata et al. (2024) muestran que la inteligencia artificial puede intervenir en diversas actividades del proceso investigativo, incluyendo la organización y revisión de literatura. Utilizada con este propósito, facilita comparar grandes volúmenes de estudios y localizar desequilibrios internos. La formulación de nuevas investigaciones surge entonces de la tensión entre lo que concentra atención y aquello que permanece insuficientemente desarrollado.

La identificación de una brecha requiere también comprobar que no se trate de un vacío producido por una búsqueda limitada. Arias y Artigas (2022) destacan la necesidad de construir problemas científicos a partir de una revisión suficientemente rigurosa de la literatura. Antes de convertir una ausencia en pregunta de investigación, conviene examinar si existen trabajos en otras disciplinas, con terminología diferente o publicados en contextos menos visibles. Una brecha aparente puede desaparecer cuando se amplía el corpus, mientras una brecha real debería mantenerse reconocible tras una exploración documental más amplia.

Las tendencias metodológicas también pueden orientar nuevas preguntas. Si un campo utiliza de forma reiterada diseños semejantes, puede resultar pertinente explorar el problema mediante

otras estrategias capaces de producir evidencia diferente. Tamayo et al. (2015) sitúan el pensamiento crítico como una capacidad necesaria para examinar problemas y argumentos, lo que resulta aplicable a la evaluación de prácticas metodológicas dominantes. Formular nuevas investigaciones no consiste únicamente en buscar temas novedosos, sino también en cuestionar cómo se ha construido la evidencia disponible y qué limitaciones produce esa reiteración.

La alfabetización científica cumple un papel relevante en esta etapa porque permite interpretar tendencias y vacíos sin depender de manera acrítica de las salidas automatizadas. Mujica (2021) destaca la importancia de desarrollar capacidades para comprender y evaluar información científica en entornos digitales. Cuando un sistema identifica un tema emergente o una zona con baja producción, el investigador necesita valorar su significado disciplinar, metodológico y conceptual. La IA puede señalar patrones, pero la decisión sobre cuáles merecen convertirse en problemas científicos exige conocimiento del campo y criterio analítico.

También es posible formular investigaciones a partir de contradicciones persistentes dentro de una tendencia consolidada. Un área puede mostrar crecimiento sostenido y, aun así, mantener resultados divergentes o marcos conceptuales poco integrados. Krenn et al. (2022) permiten considerar la inteligencia artificial como apoyo para explorar relaciones complejas entre distintos componentes del conocimiento. En estos casos, la nueva investigación no surge porque falten estudios, sino porque la evidencia existente todavía no permite explicar de manera suficiente por qué aparecen resultados incompatibles o qué condiciones determinan sus diferencias.

La pertinencia de una nueva investigación depende, además, de que la pregunta derivada del vacío sea suficientemente delimitada. Barradas (2023) analiza la inteligencia artificial como un elemento transformador de la investigación científica, contexto en el que la capacidad de procesar información puede ampliar considerablemente la detección de temas potenciales. Esa abundancia también puede generar propuestas demasiado amplias. Un problema sólido necesita precisar qué se desconoce, en qué población o contexto, mediante qué relación conceptual y por qué esa ausencia afecta la comprensión disponible del fenómeno.

La formulación basada en tendencias debe evitar otra distorsión: investigar únicamente aquello que aparece con mayor visibilidad algorítmica. Los sistemas pueden privilegiar temas con más documentos, citas o conexiones, dejando en posiciones periféricas problemas menos representados. Mata et al. (2024) permiten situar este riesgo dentro del uso más amplio de

inteligencia artificial en investigación. Una agenda científica no debería construirse solo a partir de lo que el sistema muestra como dominante. Las zonas marginales pueden contener preguntas relevantes precisamente porque han recibido menor atención o permanecen fuera de circuitos consolidados.

El uso más productivo de tendencias y vacíos consiste en convertir patrones documentales en preguntas que puedan someterse a investigación, no en tratarlos como justificaciones automáticas. Una nueva propuesta gana solidez cuando puede explicar qué limitación previa intenta resolver, qué evidencia demuestra esa limitación y qué aporte específico pretende realizar. La inteligencia artificial puede acelerar la detección de señales, pero la formulación científica comienza realmente cuando esas señales se traducen en un problema preciso, argumentado y suficientemente acotado para producir conocimiento verificable.

Capítulo

5.

Verificación inteligente de fuentes, citas y referencias

Juan Gabriel Zaruma Piña



5.1 Calidad y confiabilidad de las fuentes científicas

La calidad de una fuente científica no depende únicamente de que el documento exista, tenga una estructura académica o aparezca citado dentro de otros trabajos. Su valoración exige examinar la coherencia entre el problema planteado, el método utilizado, la evidencia presentada y las conclusiones formuladas. De la Cruz (2017) relaciona el rigor científico con la consistencia del proceso investigativo, criterio que resulta útil para evaluar documentos recuperados mediante sistemas digitales. Una fuente confiable debe ofrecer elementos suficientes para comprender cómo se construyó el conocimiento que comunica.

La confiabilidad tampoco puede reducirse al prestigio aparente de una publicación. Una referencia formalmente correcta puede contener limitaciones metodológicas importantes, mientras un documento menos visible puede ofrecer evidencia pertinente para una pregunta específica. Mugaanyi et al. (2024) muestran que los modelos de lenguaje pueden presentar problemas en la generación de citas y referencias, lo que obliga a separar dos niveles de evaluación: primero, comprobar que la fuente existe; después, valorar si su contenido posee calidad suficiente para sustentar una afirmación científica.

La inteligencia artificial complica esta tarea porque puede presentar referencias con una apariencia bibliográfica convincente. Chelli et al. (2024) documentan tasas de alucinación y problemas de exactitud en referencias generadas por ChatGPT y Bard para tareas vinculadas con revisiones sistemáticas. Esto significa que la forma de una referencia no garantiza su validez. Autor, título, revista, año y DOI deben comprobarse antes de incorporar una fuente al corpus, especialmente cuando el documento fue sugerido por una herramienta generativa y no recuperado directamente desde una base académica verificable.

La existencia de la fuente constituye solo el primer nivel de control. Una vez verificados sus datos, debe examinarse si el documento responde realmente al tema o problema que pretende respaldar. Castillo (2023) subraya la necesidad de supervisión humana cuando ChatGPT se emplea como apoyo en tareas científicas. Este criterio adquiere especial relevancia en la gestión bibliográfica, porque los sistemas pueden proponer documentos temáticamente cercanos pero insuficientes para sustentar una afirmación concreta. La pertinencia debe establecerse mediante lectura y contraste, no mediante confianza en la recomendación automática.

Otro componente de calidad corresponde a la trazabilidad del contenido utilizado. Una afirmación científica resulta más confiable cuando puede seguirse desde la interpretación realizada hasta el documento que la sustenta. Penabad et al. (2024) destacan la importancia de

transparentar el uso de inteligencia artificial en contextos científico-académicos, lo que refuerza la necesidad de documentar cómo se localizaron, procesaron y utilizaron las fuentes. La trazabilidad permite detectar errores de atribución y evita que una síntesis generada se convierta en intermediaria única entre el investigador y la evidencia original.

La consistencia interna del documento también merece atención. Un estudio puede abordar un problema relevante y, sin embargo, presentar discrepancias entre sus objetivos, métodos y conclusiones. De la Cruz (2017) permite situar el rigor en la coherencia de estas relaciones, más que en señales superficiales de formalidad académica. La evaluación de calidad requiere revisar si los resultados derivan de los procedimientos descritos y si las conclusiones respetan el alcance de la evidencia. Una fuente útil no debería sostener más de lo que su diseño permite justificar.

La integridad académica constituye otro criterio relevante para valorar la confiabilidad de una fuente y del proceso mediante el cual se incorpora a una investigación. Gallent et al. (2023) analizan los retos éticos asociados con la inteligencia artificial generativa en educación superior y la integridad académica. Aplicado a la gestión de fuentes, esto implica utilizar la IA sin ocultar su intervención, verificar el contenido recuperado y evitar atribuciones imprecisas. La calidad documental pierde valor si la fuente es manipulada, citada fuera de contexto o utilizada para respaldar afirmaciones que no contiene.

Los sistemas generativos también pueden introducir errores parciales en referencias reales. Mugaanyi et al. (2024) muestran que la confiabilidad de citas y referencias producidas por modelos de lenguaje varía y requiere comprobación. Un registro puede corresponder a un artículo existente, pero contener errores en el título, los autores o el identificador digital. Esta situación es especialmente problemática porque el error parece menor y puede pasar inadvertido. La verificación bibliográfica debe comprobar todos los elementos esenciales y no limitarse a constatar que existe una publicación de nombre parecido.

La confiabilidad debe analizarse igualmente en relación con el uso que se hará de la fuente. Un artículo puede ser adecuado para describir antecedentes y no resultar suficiente para sustentar una afirmación causal o una generalización amplia. Penabad et al. (2024) permiten reforzar la necesidad de reportar con claridad el uso de herramientas de IA y mantener criterios de responsabilidad científica. La correspondencia entre evidencia y afirmación constituye un aspecto central de calidad, porque una fuente correctamente citada puede seguir siendo utilizada de forma incorrecta si se exagera el alcance de sus resultados.

La evaluación crítica exige también prestar atención a las limitaciones reconocidas por los propios estudios. Una fuente confiable no es necesariamente aquella que presenta resultados concluyentes, sino aquella que permite comprender con claridad hasta dónde llegan sus hallazgos. Castillo (2023) destaca la importancia de mantener supervisión humana en contextos donde la inteligencia artificial interviene en la producción científica. Esa supervisión debe extenderse a la lectura de las limitaciones, porque los sistemas de síntesis pueden omitirlas y producir representaciones más categóricas que el documento original.

El control de calidad resulta especialmente importante cuando se trabaja con grandes colecciones de documentos, porque la automatización puede amplificar errores pequeños. Si una referencia incorrecta ingresa al corpus y posteriormente es resumida, clasificada y citada, el problema se propaga por varias etapas del análisis. Chelli et al. (2024) muestran que las herramientas generativas pueden introducir referencias inexactas en tareas de revisión. Por ello, la comprobación temprana de existencia, correspondencia temática y consistencia bibliográfica funciona como una barrera necesaria antes de realizar procesos analíticos posteriores.

Una fuente científica adquiere verdadero valor cuando puede ser verificada, comprendida dentro de sus límites y utilizada de acuerdo con lo que realmente demuestra. La disponibilidad de herramientas de inteligencia artificial no elimina este criterio; lo vuelve más exigente porque aumenta la velocidad con la que pueden circular referencias plausibles, resúmenes convincentes y afirmaciones descontextualizadas. La evaluación crítica debe conservar una regla básica: ninguna recomendación automática sustituye la revisión del documento original ni la responsabilidad de establecer si esa evidencia es suficiente para sustentar el argumento que se desea construir.

5.2 Verificación de existencia y correspondencia de referencias

La verificación de una referencia comienza por determinar si el documento citado existe realmente y si sus datos bibliográficos corresponden a una publicación identificable. Esta comprobación adquiere especial importancia cuando la referencia procede de una herramienta generativa, porque una estructura formalmente correcta puede ocultar información inventada o combinada de distintas fuentes. Walters y Wilder (2023) muestran que ChatGPT puede producir citas bibliográficas fabricadas y referencias reales con errores, lo que obliga a comprobar cada elemento antes de incorporarlo a un trabajo científico.

La existencia documental constituye solo el primer nivel de validación. Después debe comprobarse que autoría, título, revista, año, volumen, número, páginas e identificador digital

correspondan al mismo documento. Athaluri et al. (2023) examinan el fenómeno de las alucinaciones en referencias científicas generadas por ChatGPT y evidencian que pueden aparecer datos plausibles sin respaldo real. El problema se agrava cuando varios componentes correctos se combinan de manera errónea, porque la referencia resultante puede parecer auténtica y superar una revisión superficial.

Los identificadores persistentes facilitan la comprobación, pero tampoco deben considerarse suficientes de forma aislada. Aiumtrakul et al. (2023), al estudiar referencias obtenidas mediante ChatGPT, BingChat y Bard en búsquedas de literatura, analizaron la exactitud de las citas y la correspondencia de los DOI. Esta práctica permite distinguir entre una referencia existente y una referencia cuyos metadatos han sido alterados. Un DOI válido debe dirigir al documento citado y coincidir con sus autores, título y publicación; no basta con que el identificador exista en alguna base de datos.

La correspondencia bibliográfica implica también confirmar que la referencia recuperada sea la fuente que el sistema pretendía sugerir. Mugaanyi et al. (2024) muestran que los modelos de lenguaje presentan variaciones en precisión y confiabilidad al producir citas y referencias académicas. Un error aparentemente menor, como modificar el año o sustituir parte del título, puede conducir a otro trabajo o dificultar su recuperación. La verificación debe realizarse como una comparación sistemática entre la referencia generada y los metadatos del documento original, no mediante una revisión visual aproximada.

Chelli et al. (2024) muestran que las herramientas generativas pueden producir tanto referencias válidas como referencias alucinadas en tareas vinculadas con revisiones sistemáticas. Esta coexistencia dificulta la detección, porque la presencia de varias citas correctas puede aumentar la confianza en el resto de la lista. Por ello, la validación no debería hacerse por muestreo cuando las referencias serán utilizadas como evidencia. Cada registro necesita comprobarse de manera independiente, especialmente si fue incorporado mediante una respuesta generada automáticamente y no recuperado directamente desde una fuente bibliográfica verificable.

Tabla 13
Controles para verificar existencia y correspondencia de referencias científicas

Nivel de comprobación	Elemento revisado	Criterio de correspondencia	Riesgo si no se verifica
Existencia	Documento citado	La publicación puede localizarse de forma independiente	Incorporar referencias fabricadas

Nivel de comprobación	Elemento revisado	Criterio de correspondencia	Riesgo si no se verifica
Autoría	Autores y orden	Coinciden con el documento original	Atribuir trabajos a autores incorrectos
Datos editoriales	Revista, año, volumen, número y páginas	Pertenecen a la misma publicación	Combinar metadatos de documentos diferentes
Título	Denominación completa del trabajo	Coincide con la fuente recuperada	Confundir publicaciones semejantes
DOI	Identificador digital	Resuelve hacia el documento citado	Asociar un DOI real con una referencia distinta
Correspondencia temática	Contenido de la fuente	El documento guarda relación con el uso previsto	Citar una fuente existente pero inadecuada

Nota. La tabla distingue la comprobación de existencia de la verificación integral de correspondencia bibliográfica y temática.

La tabla permite observar que una referencia puede superar un nivel de control y fallar en otro. Walters y Wilder (2023) diferencian precisamente entre fabricación completa y errores bibliográficos dentro de referencias vinculadas con documentos existentes. Esta distinción es metodológicamente importante porque los errores parciales suelen ser menos visibles. Una publicación real con título modificado o autores incompletos puede incorporarse al manuscrito sin levantar sospechas, aunque después dificulte la trazabilidad y genere inconsistencias en las citas y en la lista final de referencias.

La verificación debe extenderse finalmente al contenido, porque una referencia real y bibliográficamente correcta puede no corresponder con la afirmación para la que fue recuperada. Aiumtrakul et al. (2023) muestran que la evaluación de herramientas de IA para búsquedas bibliográficas requiere considerar no solo la existencia de las citas, sino también su precisión. La correspondencia temática exige revisar el artículo y determinar qué estudio realmente. Esta comprobación evita utilizar una fuente válida para sustentar un argumento que excede o simplemente no aparece en su contenido.

Un procedimiento riguroso de verificación debe tratar cada referencia como un conjunto de datos que necesita concordancia interna. Autoría, título, publicación, fecha e identificador deben conducir al mismo documento y ese documento debe responder al uso científico que se pretende darle. La automatización puede ayudar a detectar inconsistencias, pero no debería transformar la validación en un acto mecánico de aceptación. El criterio decisivo es que

cualquier lector pueda localizar la fuente, confirmar sus metadatos y comprobar que la evidencia atribuida procede efectivamente de ella.

5.3 Comprobación automatizada de autores, títulos, revistas y DOI

La comprobación automatizada de metadatos bibliográficos busca verificar que los elementos principales de una referencia correspondan al mismo documento científico. Autores, título, revista, año y DOI constituyen piezas que deben mantener coherencia entre sí. Walters y Wilder (2023) muestran que los modelos generativos pueden producir referencias completamente fabricadas y, también, referencias reales con errores parciales. Esta diferencia vuelve necesaria revisar cada componente de manera independiente, porque una referencia puede parecer correcta a simple vista y contener una combinación de datos procedentes de publicaciones distintas.

La autoría requiere un control específico, ya que los errores pueden afectar nombres, iniciales, orden de aparición o número de autores. Mugaanyi et al. (2024) muestran que la confiabilidad de las citas y referencias generadas por grandes modelos de lenguaje presenta variaciones entre disciplinas. Un sistema automatizado de verificación debe contrastar la autoría con los metadatos de la fuente original y detectar discrepancias antes de aceptar el registro. La coincidencia parcial no resulta suficiente cuando la atribución correcta constituye un elemento esencial de la trazabilidad académica.

El título representa otro punto crítico porque puede sufrir modificaciones sutiles que dificultan la recuperación del documento. Athaluri et al. (2023) examinan referencias alucinadas generadas por ChatGPT y muestran cómo pueden aparecer formulaciones plausibles que no corresponden a publicaciones reales. La comprobación automatizada puede identificar diferencias entre el título proporcionado y el registrado oficialmente, pero debe considerar variaciones legítimas de puntuación o capitalización. El objetivo no es exigir coincidencia tipográfica absoluta, sino verificar que la referencia conduce inequívocamente al mismo trabajo científico.

La revista también debe comprobarse como parte del conjunto bibliográfico. Un sistema generativo puede asociar un título real con una publicación incorrecta o utilizar el nombre de una revista plausible que no corresponde al artículo citado. Chelli et al. (2024) muestran que la precisión de las referencias producidas por herramientas generativas puede variar de manera considerable en tareas vinculadas con revisiones sistemáticas. La validación automática debe

contrastar revista, volumen, número y año para reducir el riesgo de combinar metadatos correctos de fuentes diferentes.

El DOI ofrece una ventaja importante porque funciona como identificador persistente, pero no debe interpretarse como garantía automática de exactitud. Aiumtrakul et al. (2023) comprobaron la correspondencia de DOI en referencias producidas por herramientas de IA y encontraron la necesidad de verificar si el identificador realmente conduce al documento citado. Un DOI válido puede pertenecer a otra publicación. Por ello, la comprobación automatizada debe resolver el identificador y contrastar los metadatos recuperados con la referencia completa antes de considerarla bibliográficamente correcta.

La automatización de este proceso puede organizarse como una secuencia de contrastes. Primero se comprueba la existencia del DOI o del registro; después se cotejan autores, título, revista y fecha. Mugaanyi et al. (2024) muestran que la evaluación de referencias generadas por modelos de lenguaje exige revisar varios componentes simultáneamente. Este enfoque reduce falsos positivos, porque una coincidencia en el título no compensa una autoría incorrecta, del mismo modo que un DOI existente no valida una referencia si corresponde a otro documento.

Los errores parciales son especialmente problemáticos porque suelen superar revisiones superficiales. Walters y Wilder (2023) distinguen entre referencias fabricadas y referencias vinculadas con publicaciones reales pero afectadas por errores bibliográficos. Esta segunda categoría puede ser más difícil de detectar, ya que el lector reconoce algunos elementos y asume que el resto también es correcto. La comprobación automatizada debe señalar discrepancias mínimas y permitir revisar su importancia, especialmente cuando afectan identificación de autores, título completo o correspondencia del DOI con la fuente citada.

La validación también debe contemplar casos en los que no existe DOI. La ausencia de este identificador no convierte una fuente en inválida, pero obliga a utilizar otros elementos de correspondencia. Athaluri et al. (2023) muestran que las alucinaciones bibliográficas pueden producir combinaciones convincentes de autor, título y revista, por lo que la verificación debe apoyarse en múltiples campos. Cuando no existe un identificador persistente, la coincidencia entre autoría, título, publicación y fecha adquiere mayor importancia para establecer si el documento puede localizarse de forma independiente.

Otro problema aparece cuando las herramientas de IA modifican una referencia real al intentar normalizarla. Chelli et al. (2024) muestran que los sistemas generativos pueden presentar

problemas de exactitud incluso cuando trabajan con literatura científica. Una corrección automática de estilo puede alterar accidentalmente el título, omitir autores o introducir datos no presentes en el original. Por esa razón, los procedimientos de verificación deben realizarse después de cualquier transformación bibliográfica y no asumir que una referencia previamente correcta seguirá siéndolo una vez procesada por un sistema generativo.

Figura 9
Elementos de comprobación bibliográfica



Nota. La figura reúne los elementos básicos que deben contrastarse para verificar la correspondencia y exactitud de una referencia bibliográfica.

La comprobación automatizada resulta más sólida cuando conserva evidencia de cada coincidencia y discrepancia detectada. Aiumtrakul et al. (2023) muestran la importancia de evaluar la precisión de las referencias recuperadas mediante herramientas de IA. Aplicado a un flujo de verificación, esto implica registrar qué campos coincidieron, cuáles fueron corregidos y qué fuente se utilizó para validar el registro. Esta trazabilidad permite revisar posteriormente las modificaciones y evita que la corrección automática se convierta en un proceso opaco difícil de auditar.

El propósito de automatizar la comprobación no debe ser sustituir la validación bibliográfica, sino reducir errores repetitivos y hacer visibles inconsistencias que podrían pasar inadvertidas. Un sistema útil identifica discrepancias, propone correspondencias y facilita volver a la fuente original, pero no debería decidir sin control cuándo dos registros representan el mismo documento. La precisión académica exige que cada referencia pueda reconstruirse como una unidad coherente. Solo cuando autores, título, revista, fecha e identificador apuntan al mismo trabajo puede considerarse que el registro ha sido verificado de manera suficiente.

5.4 Evaluación de la correspondencia entre afirmaciones y citas

La correspondencia entre una afirmación y la cita que la acompaña constituye una condición básica de la argumentación científica. Una referencia puede existir, estar correctamente escrita y pertenecer al tema general del manuscrito, pero aun así no respaldar la proposición concreta que se le atribuye. Wadden y Lo (2021) muestran que la verificación de afirmaciones científicas exige relacionar una proposición con evidencia capaz de apoyarla, refutarla o resultar insuficiente. Esta distinción permite evaluar las citas por su función probatoria y no solo por su presencia formal.

La primera dificultad consiste en delimitar con precisión qué afirma realmente una oración. Una misma frase puede contener varias proposiciones sobre frecuencia, causalidad, asociación o alcance, mientras la fuente citada respalda solo una de ellas. Wadden et al. (2022) abordan la verificación científica mediante la relación entre afirmaciones y evidencia recuperada en colecciones amplias. Este enfoque muestra que comprobar una cita requiere descomponer el argumento cuando sea necesario y determinar qué fragmento documental sostiene cada componente, evitando atribuir a una única referencia un respaldo más amplio del que realmente ofrece.

El contenido de la fuente debe examinarse en su contexto y no mediante frases aisladas. Un resultado puede parecer favorable a una afirmación cuando se extrae de forma independiente, pero adquirir un significado diferente al considerar la población, el método o las limitaciones del estudio. Tang et al. (2023) muestran que la síntesis de evidencia mediante modelos de lenguaje requiere preservar información contextual suficiente para mantener fidelidad con las fuentes. Esta exigencia se aplica igualmente a las citas: el respaldo depende de cómo fue obtenido e interpretado el hallazgo citado.

Una relación de apoyo tampoco debe confundirse con una coincidencia temática. Dos documentos pueden abordar el mismo fenómeno sin responder a la misma pregunta ni sostener conclusiones equivalentes. Wadden y Lo (2021) permiten diferenciar entre evidencia que respalda una afirmación y evidencia que simplemente guarda relación con ella. Esta separación es especialmente importante en búsquedas automatizadas, donde la proximidad semántica puede conducir a seleccionar artículos cercanos al tema. La correspondencia exige un vínculo más estricto entre la proposición escrita y el contenido verificable de la fuente.

La inteligencia artificial puede apoyar esta evaluación mediante sistemas que recuperan evidencia y clasifican su relación con una afirmación. Wadden et al. (2022) muestran que la

verificación científica en dominio abierto puede combinar recuperación documental con identificación de evidencia pertinente. Esta capacidad permite revisar grandes volúmenes de literatura y localizar fragmentos potencialmente favorables o contrarios a una proposición. Sin embargo, la clasificación automática debe interpretarse como una ayuda para la comprobación, porque la relación final puede depender de matices metodológicos que no siempre quedan reflejados en una etiqueta.

Las afirmaciones generales requieren un control especialmente estricto. Una fuente puede demostrar una relación en una muestra delimitada, mientras el texto la presenta como válida para poblaciones o contextos más amplios. Krenn et al. (2022) plantean que la inteligencia artificial puede contribuir a organizar conocimiento científico complejo, pero la comprensión depende de conservar las condiciones que delimitan cada resultado. La correspondencia entre cita y afirmación se debilita cuando el argumento amplía población, causalidad, intensidad o generalidad más allá de lo que la evidencia original permite sostener.

También pueden producirse errores cuando una cita se utiliza para atribuir causalidad a un estudio que únicamente reporta asociación. La semejanza verbal entre ambos tipos de afirmación puede ocultar una diferencia metodológica importante. Tang et al. (2023) evidencian que la síntesis automatizada debe conservar el significado de los resultados y evitar reformulaciones que alteren su alcance. Al evaluar una cita, conviene identificar qué tipo de relación reportó realmente el documento y comprobar que el verbo empleado en el manuscrito no convierta una observación limitada en una conclusión más fuerte.

Las citas pueden cumplir además funciones distintas dentro de una argumentación. Algunas ofrecen evidencia directa, otras contextualizan un problema y otras presentan enfoques con los que el texto discute o discrepa. Gilbert et al. (2024) destacan la utilidad de modelos aumentados que mantienen las respuestas vinculadas con información verificable. Este principio permite exigir que cada cita conserve una función reconocible dentro del argumento. Si una referencia aparece únicamente por proximidad temática, sin que pueda identificarse qué aporta a la afirmación concreta, su inclusión pierde valor justificativo.

La evaluación debe considerar igualmente las contradicciones entre la formulación del autor citado y la interpretación introducida por quien escribe. Wadden et al. (2022) muestran que una misma afirmación puede encontrar evidencia de apoyo o refutación dentro de la literatura científica. Ignorar esta diversidad y seleccionar solo documentos favorables puede producir una argumentación aparentemente sólida pero incompleta. Cuando existen resultados

divergentes, la correspondencia exige representar esa situación con precisión y evitar presentar como consenso aquello que la evidencia disponible mantiene todavía bajo discusión o condicionado por circunstancias específicas.

Los sistemas de curación de información pueden reducir parte de estos problemas cuando permiten rastrear la procedencia de cada afirmación. Gilbert et al. (2024) analizan modelos diseñados para limitar alucinaciones mediante vínculos con información recuperada. En la evaluación de citas, esta trazabilidad facilita volver desde una afirmación hasta el pasaje documental que supuestamente la sustenta. El procedimiento resulta más confiable que aceptar una referencia completa sin revisar su contenido, porque obliga a comprobar la relación directa entre el argumento formulado y la evidencia efectivamente localizada.

La verificación también debe distinguir entre evidencia insuficiente y evidencia contradictoria. La ausencia de respaldo directo en una fuente no significa necesariamente que el documento refute la afirmación. Wadden y Lo (2021) muestran que la clasificación de evidencia científica requiere reconocer situaciones donde la información disponible no permite confirmar ni negar una proposición. Esta categoría resulta importante para evitar interpretaciones forzadas. Una cita científicamente adecuada debe aportar evidencia pertinente; cuando la fuente permanece neutral frente a la afirmación, su uso como respaldo puede crear una certeza que el documento no ofrece.

La prueba decisiva consiste en formular una pregunta sencilla frente a cada cita: ¿qué parte exacta de esta afirmación puede comprobarse en la fuente indicada? Si la respuesta exige añadir supuestos, ampliar resultados o reinterpretar el alcance del estudio, la correspondencia necesita revisión. La inteligencia artificial puede facilitar la localización de fragmentos y detectar relaciones posibles, pero no convierte automáticamente una fuente pertinente en evidencia suficiente. Mantener este control evita que las citas funcionen como adornos de autoridad y conserva su función esencial: hacer verificable la cadena que conecta argumento y evidencia científica.

5.5 Identificación de referencias falsas, incompletas o inconsistentes

La identificación de referencias falsas, incompletas o inconsistentes constituye una tarea necesaria cuando se utilizan sistemas generativos para apoyar búsquedas, redacción o gestión bibliográfica. Walters y Wilder (2023) muestran que ChatGPT puede producir tanto referencias fabricadas como registros asociados con publicaciones reales pero afectados por errores. Esta distinción es importante porque no todos los problemas bibliográficos presentan la misma

gravedad. Una referencia inexistente exige ser descartada, mientras una referencia incompleta o inconsistente puede requerir reconstrucción y verificación antes de ser utilizada en un trabajo científico.

Las referencias falsas suelen presentar una combinación plausible de autores, títulos y revistas que imita la estructura de una publicación académica. Athaluri et al. (2023) examinan este fenómeno en referencias generadas por ChatGPT y muestran cómo la alucinación puede producir datos que parecen verosímiles sin corresponder a una fuente real. El riesgo aumenta cuando el usuario reconoce alguno de los elementos o cuando el título resulta coherente con el tema buscado. La apariencia académica, por tanto, no constituye evidencia de existencia documental.

Una referencia incompleta presenta un problema distinto. Puede corresponder a un documento real, pero omitir autores, páginas, volumen, número, año o identificador digital. Mugaanyi et al. (2024) muestran que los modelos de lenguaje pueden producir citas y referencias con distintos niveles de precisión bibliográfica. Estas omisiones afectan la trazabilidad, porque dificultan localizar la fuente y comprobar su contenido. La corrección exige contrastar el registro con la publicación original y completar únicamente los datos verificables, evitando reconstruir campos faltantes mediante inferencias no sustentadas.

Las inconsistencias aparecen cuando los distintos elementos de la referencia no corresponden entre sí. Un título real puede asociarse con una revista equivocada, un DOI válido puede pertenecer a otro artículo o los autores pueden proceder de una publicación diferente. Chelli et al. (2024) documentan problemas de precisión en referencias producidas por herramientas generativas en tareas de revisión sistemática. Estos errores son especialmente difíciles de detectar porque cada componente puede parecer correcto por separado. La validación debe comprobar la coherencia del registro como una unidad bibliográfica completa.

La identificación de problemas también requiere distinguir errores formales de errores sustantivos. Una referencia puede presentar una puntuación incorrecta y seguir identificando claramente el documento, mientras otra puede tener una apariencia estilística perfecta y corresponder a una fuente inexistente. Aiumtrakul et al. (2023) muestran la necesidad de evaluar la exactitud de referencias generadas por herramientas de IA, incluyendo su correspondencia con identificadores y publicaciones reales. La prioridad debe situarse en la existencia y coherencia bibliográfica antes de corregir aspectos secundarios de estilo o formato.

Tabla 14
Tipos de errores en referencias generadas o procesadas con inteligencia artificial

Tipo de problema	Señal de alerta	Comprobación necesaria	Acción recomendada
Referencia falsa	No aparece en fuentes verificables	Buscar autor, título y DOI de forma independiente	Descartar si no puede confirmarse
Referencia incompleta	Faltan campos bibliográficos esenciales	Contrastar con el documento original	Completar solo datos verificados
Referencia inconsistente	Los campos conducen a documentos distintos	Revisar cada elemento y su correspondencia	Reconstruir el registro desde la fuente real
DOI incorrecto	El identificador dirige a otra publicación	Resolver el DOI y comparar metadatos	Sustituir por el identificador correcto
Autoría alterada	Nombres u orden no coinciden	Contrastar con la publicación original	Corregir la atribución
Título modificado	Existen diferencias sustantivas en el nombre	Comparar con el título oficial	Restaurar la denominación original

Nota. La tabla diferencia errores bibliográficos que pueden parecer similares, pero requieren procedimientos de comprobación distintos.

La clasificación de errores ayuda a evitar correcciones automáticas indiscriminadas. Walters y Wilder (2023) muestran que la fabricación completa y los errores parciales constituyen fenómenos distintos, por lo que no deben tratarse de la misma forma. Un sistema puede detectar una coincidencia aproximada y proponer una sustitución, pero esa operación puede vincular accidentalmente una referencia falsa con un artículo real de título semejante. La corrección debe partir de evidencia bibliográfica suficiente y no de la necesidad de “hacer coincidir” un registro con alguna publicación existente.

Los errores de DOI merecen un control específico porque suelen producir una falsa sensación de seguridad. Athaluri et al. (2023) muestran que las referencias alucinadas pueden incluir datos bibliográficos convincentes, mientras Aiumtrakul et al. (2023) evidencian la importancia de comprobar los identificadores asociados con las citas recuperadas. Un DOI debe considerarse válido únicamente cuando conduce al mismo documento descrito por el resto de la referencia. La mera existencia del identificador no confirma la exactitud del registro si los demás metadatos pertenecen a otra publicación.

La revisión automatizada puede utilizar reglas de consistencia para señalar referencias sospechosas, pero la decisión final requiere contrastar la fuente. Mugaanyi et al. (2024) muestran que la confiabilidad bibliográfica de los modelos de lenguaje no es uniforme, lo que hace necesario mantener controles independientes. Una referencia marcada como dudosa debe

revisarse desde sus metadatos originales, mientras una referencia aceptada automáticamente también puede requerir verificación cuando será utilizada como evidencia central. La detección funciona mejor como mecanismo de alerta que como sustituto de la comprobación documental.

Un sistema de referencias confiable se construye descartando lo que no puede verificarse y corrigiendo únicamente aquello que puede reconstruirse desde una fuente real. Forzar una coincidencia para conservar una cita introduce más riesgo que reconocer que el registro no es válido. La coherencia bibliográfica exige que cada componente pueda comprobarse y que todos conduzcan al mismo documento. Esta disciplina reduce la propagación de errores y evita que una referencia plausible, pero defectuosa, termine adquiriendo legitimidad por repetición dentro de nuevos textos científicos.

5.6 Trazabilidad de la evidencia científica utilizada

La trazabilidad de la evidencia científica permite reconstruir el recorrido que conecta una afirmación, los datos que la respaldan y la fuente documental de la que proceden. Este principio adquiere mayor relevancia cuando la inteligencia artificial interviene en procesos de búsqueda, extracción o síntesis, porque parte de las operaciones puede quedar oculta para el usuario. Penabad et al. (2024) destacan la necesidad de reportar de manera transparente el empleo de estas herramientas, lo que exige conservar información suficiente para identificar cómo se obtuvo y utilizó cada evidencia.

La trazabilidad comienza con una identificación inequívoca de las fuentes incorporadas al corpus. Cada documento debe conservar sus metadatos esenciales y mantenerse vinculado con las decisiones que justificaron su inclusión. Tapia y Recalde (2024) muestran la importancia de organizar de manera sistemática los procesos de gestión documental, criterio que resulta aplicable a la construcción de colecciones científicas. Cuando los registros se desordenan, duplican o separan de sus datos originales, aumenta la dificultad para determinar de dónde procede una afirmación y cómo fue incorporada al análisis.

El seguimiento debe mantenerse también cuando la información pasa del documento original a matrices, resúmenes o estructuras intermedias. Wadden et al. (2022) muestran que la verificación científica puede relacionar afirmaciones concretas con evidencia recuperada de publicaciones académicas. Este principio permite exigir que cada dato extraído conserve un vínculo con el fragmento que lo sustenta. Una matriz de resultados pierde valor si sus celdas contienen información que ya no puede rastrearse hasta la fuente. La eficiencia analítica depende, por tanto, de preservar esa conexión documental.

La inteligencia artificial puede facilitar la extracción de grandes cantidades de información, pero también puede introducir transformaciones que dificulten identificar el origen exacto de una formulación. Penabad et al. (2024) refuerzan la importancia de declarar cómo intervino la IA dentro del proceso científico. Si una herramienta resumió, clasificó o reformuló evidencia, esa operación debe diferenciarse de la información obtenida directamente del artículo. La trazabilidad exige saber qué parte proviene de la fuente, qué parte corresponde al procesamiento automático y qué interpretación fue incorporada posteriormente por el investigador.

La reproducibilidad guarda una relación estrecha con esta exigencia. Semmelrock et al. (2025) analizan barreras y factores que afectan la reproducibilidad en investigaciones basadas en aprendizaje automático, donde la documentación de procedimientos resulta fundamental. En el manejo automatizado de literatura científica, registrar herramientas, decisiones y transformaciones permite reconstruir cómo se llegó desde un corpus inicial hasta una síntesis determinada. La trazabilidad no busca solo identificar fuentes, sino conservar la secuencia de operaciones que modificaron, agruparon o priorizaron la evidencia durante el proceso de análisis.

Tabla 15
Trazabilidad de la evidencia en procesos científicos asistidos por inteligencia artificial

Etapas	Elemento que debe conservarse	Función de trazabilidad	Riesgo si se pierde
Recuperación	Fuente, consulta y fecha de localización	Reconstruir el origen del documento	Desconocer cómo ingresó al corpus
Selección	Criterio de inclusión o exclusión	Justificar decisiones documentales	Construir un corpus difícil de auditar
Extracción	Dato y fragmento de origen	Vincular información con evidencia directa	Separar el dato de su contexto
Procesamiento con IA	Herramienta y operación realizada	Diferenciar fuente y transformación automatizada	Confundir generación con evidencia
Síntesis	Afirmaciones y documentos de respaldo	Comprobar la construcción del argumento	Presentar conclusiones sin procedencia verificable
Revisión	Cambios y correcciones efectuadas	Mantener historial del análisis	Perder decisiones que modificaron los resultados

Nota. La tabla organiza los puntos de control necesarios para reconstruir el recorrido de la evidencia desde su recuperación hasta su incorporación en una síntesis científica.

La secuencia presentada muestra que la trazabilidad no puede añadirse únicamente al final del proceso. Loza et al. (2026) estudian la reproducibilidad en flujos de trabajo bioinformáticos y permiten comprender la importancia de estructurar procedimientos desde sus primeras etapas. Cuando las decisiones se documentan solo después de completar el análisis, pueden perderse cambios, exclusiones o transformaciones relevantes. En cambio, un registro continuo facilita identificar en qué momento se modificó la información y permite revisar errores sin reconstruir retrospectivamente todo el procedimiento seguido.

Los grafos de conocimiento pueden contribuir a esta tarea cuando representan explícitamente relaciones entre entidades y fuentes. Brack et al. (2022) muestran que la construcción de grafos abiertos de investigación depende de requisitos de calidad y estrategias de estructuración suficientemente claras. Estas representaciones pueden facilitar el seguimiento entre documentos, conceptos y afirmaciones, siempre que las relaciones mantengan información sobre su procedencia. Un vínculo visual resulta insuficiente si no puede establecerse qué documento lo originó o qué criterio justificó su incorporación dentro de la estructura científica.

La trazabilidad también protege frente a problemas derivados de procedimientos automatizados difíciles de reproducir. Kapoor y Narayanan (2023) muestran que la investigación basada en aprendizaje automático puede verse afectada por fallas metodológicas que comprometen la reproducibilidad. En el análisis documental, una clasificación o síntesis generada por IA debe poder examinarse mediante sus entradas y decisiones principales. Si solo se conserva el resultado final, resulta imposible valorar cuánto depende de la evidencia original y cuánto de particularidades del modelo o de la configuración utilizada durante el procesamiento.

El seguimiento de la evidencia permite además detectar cuándo una afirmación ha sufrido transformaciones sucesivas. Un resultado puede pasar del artículo original a una extracción automática, luego a una matriz y finalmente a una síntesis narrativa. Cada paso introduce la posibilidad de simplificación o pérdida de contexto. Wadden et al. (2022) muestran la utilidad de vincular afirmaciones con evidencia verificable. Mantener esa relación durante todas las fases evita que una interpretación secundaria termine presentándose como si fuera un hallazgo directamente reportado por la fuente científica consultada.

Una cadena de evidencia sólida debe permitir avanzar y retroceder entre argumento, dato y documento sin perder información esencial. Este criterio convierte la trazabilidad en una propiedad del proceso completo y no en un requisito exclusivamente bibliográfico. Cuando cada transformación queda registrada, es posible revisar decisiones, corregir errores y discutir

el alcance de una afirmación con base en su procedencia real. La inteligencia artificial puede intervenir en varias etapas, pero su uso es científicamente más seguro cuando ninguna operación rompe la posibilidad de volver a la evidencia original.

5.7 Uso responsable de inteligencia artificial en la gestión de citas y referencias

El uso responsable de inteligencia artificial en la gestión de citas y referencias exige diferenciar entre asistencia técnica y validación académica. Una herramienta puede localizar registros, completar datos, sugerir formatos o detectar inconsistencias, pero estas funciones no eliminan la responsabilidad de comprobar la fuente original. Penabad et al. (2024) subrayan la necesidad de transparentar el uso de inteligencia artificial en contextos científico-académicos. Esta exigencia resulta especialmente pertinente cuando la herramienta interviene en decisiones bibliográficas que afectan la trazabilidad, la atribución de autoría y la verificabilidad del texto científico.

La supervisión humana constituye un principio central porque los sistemas generativos pueden producir referencias convincentes sin garantizar su exactitud. Castillo (2023) destaca la importancia del control humano cuando ChatGPT se utiliza como apoyo en la escritura científica. Aplicado a la gestión bibliográfica, esto implica revisar cada referencia antes de incorporarla al manuscrito, especialmente cuando fue generada, completada o corregida mediante IA. La automatización puede reducir tareas repetitivas, pero no debería convertir la aceptación de citas en un proceso basado únicamente en la confianza en el sistema.

El riesgo de fabricación bibliográfica obliga a establecer controles adicionales. Walters y Wilder (2023) muestran que ChatGPT puede generar referencias inexistentes y también introducir errores en registros correspondientes a publicaciones reales. Esta combinación dificulta la detección, porque una lista puede contener citas correctas junto con otras parcialmente falsas. El uso responsable requiere verificar existencia, autoría, título, revista y DOI de manera independiente. La presencia de algunos registros válidos no debe utilizarse como argumento para asumir que toda la bibliografía generada por el sistema es confiable.

La fiabilidad tampoco es uniforme entre tareas y disciplinas. Mugaanyi et al. (2024) muestran variaciones en el desempeño de modelos de lenguaje al producir citas y referencias en distintos campos académicos. Esta evidencia desaconseja tratar la IA como una fuente bibliográfica autosuficiente. Su utilidad puede concentrarse en tareas de apoyo, como detectar campos faltantes o reorganizar referencias, mientras la comprobación debe realizarse sobre registros

verificables. El nivel de confianza otorgado al sistema debería depender de la operación concreta y no de una valoración general de la herramienta.

La responsabilidad también incluye distinguir entre una fuente sugerida por IA y una fuente efectivamente consultada. Ros y Pérez (2023) analizan el papel de ChatGPT en la escritura científica y advierten sobre sus límites como herramienta de apoyo. Citar un documento solo porque aparece mencionado en una respuesta generada implica asumir información que puede no haber sido revisada. La práctica adecuada exige acceder a la publicación, comprobar sus metadatos y verificar que su contenido respalde la afirmación correspondiente antes de incorporarla a la lista de referencias.

La integridad académica se ve comprometida cuando la IA altera la relación entre autor, evidencia y cita. Gallent et al. (2023) analizan los desafíos que la inteligencia artificial generativa plantea para la ética y la integridad académica en educación superior. En la gestión de referencias, esta preocupación se traduce en evitar citas decorativas, fuentes no consultadas o atribuciones construidas automáticamente. Una referencia debe cumplir una función demostrable dentro del argumento. Su inclusión no debería responder al deseo de aumentar el número de fuentes ni a una recomendación algorítmica sin verificación.

El uso responsable requiere además registrar cuándo la inteligencia artificial intervino de forma sustantiva en el tratamiento bibliográfico. Penabad et al. (2024) proponen criterios para el reporte de IA en publicaciones científico-académicas, lo que refuerza la necesidad de transparencia. Si una herramienta fue utilizada para depurar referencias, identificar inconsistencias o apoyar búsquedas, conviene que el proceso metodológico conserve evidencia suficiente de esa intervención. La transparencia permite distinguir entre decisiones tomadas por el investigador y operaciones automatizadas que pudieron influir en la selección o presentación de las fuentes.

La dimensión ética también involucra el tipo de confianza que se deposita en los sistemas generativos. Contreras et al. (2024) examinan desafíos éticos asociados con el uso de inteligencia artificial en investigación científica. En el manejo de citas, confiar de forma acrítica en una salida generada puede transferir errores al manuscrito y afectar la calidad de la evidencia presentada. La responsabilidad no consiste en rechazar la automatización, sino en definir límites claros: la IA puede asistir en tareas operativas, pero las decisiones que afectan atribución y sustento científico requieren comprobación humana.

Figura 10

¿Cómo usar la IA de manera responsable en la gestión de citas y referencias?



Nota. La figura sintetiza criterios de control que orientan un uso responsable de la inteligencia artificial en la gestión y verificación de citas y referencias.

El control debe mantenerse incluso cuando la herramienta presenta un nivel alto de acierto. Mugaanyi et al. (2024) muestran que el desempeño de los modelos en citas y referencias puede ser útil, pero no completamente confiable. Esto implica que la verificación no puede reservarse únicamente para registros que parezcan sospechosos. Las referencias más plausibles también deben comprobarse, porque los errores parciales suelen ser precisamente los menos visibles. Una rutina consistente de validación reduce el riesgo de incorporar información defectuosa por exceso de confianza en la fluidez o apariencia académica del sistema.

La gestión responsable de citas y referencias parte de una regla sencilla: ninguna facilidad tecnológica debe romper la cadena entre afirmación, fuente y verificación. La inteligencia artificial puede ayudar a localizar, ordenar y revisar registros, pero el control final requiere comprobar el documento y su correspondencia con el argumento. Esta práctica mantiene la utilidad de la automatización sin delegar en ella decisiones que afectan la integridad científica. El criterio no es cuánto trabajo puede hacer el sistema, sino qué tareas pueden automatizarse sin perder trazabilidad, exactitud y responsabilidad académica.

Conclusión

La incorporación de inteligencia artificial a la búsqueda y análisis de información científica modifica de manera sustantiva las tareas documentales vinculadas con la investigación. La localización de literatura ya no depende exclusivamente de palabras clave y operadores tradicionales, sino que puede apoyarse en recuperación semántica, recomendación automatizada, procesamiento del lenguaje natural y sistemas capaces de establecer relaciones contextuales entre documentos. Esta ampliación mejora las posibilidades de exploración, pero también exige formular preguntas precisas, establecer criterios de selección y conservar el control metodológico sobre los resultados recuperados.

La utilidad de estas tecnologías aumenta cuando la búsqueda se concibe como un proceso estructurado y reproducible. Formular consultas, refinarlas, filtrar documentos y priorizar resultados son operaciones diferentes que requieren criterios explícitos. La automatización puede reducir tareas repetitivas y facilitar el tratamiento de grandes volúmenes documentales, pero una respuesta rápida o una elevada cantidad de resultados no garantizan pertinencia científica. La calidad del proceso depende de que pueda explicarse cómo se construyó el corpus, qué decisiones determinaron las exclusiones y qué intervención tuvo la inteligencia artificial en cada etapa.

Los mapas digitales del conocimiento amplían este análisis al representar relaciones entre autores, documentos, conceptos y áreas científicas. Redes de coautoría, cocitación, acoplamiento bibliográfico, grafos de conocimiento y modelos temáticos permiten observar configuraciones que difícilmente podrían reconocerse mediante una lectura lineal. Sin embargo, estos mapas no constituyen reproducciones neutrales del conocimiento.

Cada representación depende del corpus utilizado, los metadatos disponibles, los criterios de agrupamiento y las relaciones seleccionadas. Por ello, la centralidad, proximidad o pertenencia a un clúster deben interpretarse como propiedades de una representación y no como juicios automáticos de relevancia. La lectura automatizada introduce otra transformación al permitir extraer información, clasificar documentos, generar resúmenes y comparar simultáneamente numerosas investigaciones. Estas capacidades pueden facilitar la identificación de objetivos, métodos, resultados, limitaciones y patrones recurrentes dentro de colecciones extensas.

El problema aparece cuando la condensación elimina condiciones necesarias para interpretar la evidencia o cuando una salida generada introduce información no presente en las fuentes. La lectura automatizada resulta metodológicamente útil cuando reduce el trabajo de

organización y exploración sin romper la relación entre cada dato procesado y el documento del que procede. La detección automatizada de tendencias, temas emergentes, vacíos y contradicciones amplía las posibilidades de estudiar la estructura y evolución de un campo científico.

Su valor reside en reconocer patrones que orienten preguntas más específicas, no en producir diagnósticos definitivos mediante frecuencias o agrupamientos. Un aumento de publicaciones no demuestra por sí mismo la consolidación de una línea, del mismo modo que una baja presencia documental no constituye automáticamente una brecha de conocimiento. Las señales algorítmicas requieren interpretación temporal, metodológica y conceptual antes de utilizarse para justificar nuevas investigaciones.

La verificación de fuentes adquiere especial relevancia porque los sistemas generativos pueden producir referencias inexistentes, datos bibliográficos incorrectos o citas que no respaldan las afirmaciones para las que fueron utilizadas. Comprobar autores, títulos, revistas, fechas y DOI constituye únicamente una primera etapa. También debe verificarse que la publicación realmente sostenga el argumento atribuido y que sus resultados no hayan sido ampliados fuera de sus condiciones originales.

La trazabilidad permite mantener esta relación entre afirmación, evidencia y fuente, incluso cuando la información atraviesa procesos sucesivos de extracción, clasificación, síntesis o reformulación automatizada. El criterio que articula estas nuevas formas de buscar y analizar información científica es, por tanto, la combinación entre capacidad computacional y control crítico.

La inteligencia artificial puede ampliar el volumen de literatura que es posible explorar, descubrir conexiones poco visibles y reducir tareas documentales que consumen tiempo, pero no determina por sí misma qué evidencia es suficiente, qué relación resulta científicamente significativa o qué vacío merece convertirse en un problema de investigación. Su incorporación resulta más rigurosa cuando cada operación automatizada permanece verificable, explicable y subordinada a decisiones metodológicas explícitas.

La transformación más relevante no consiste en reemplazar la búsqueda, la lectura o el análisis realizados por el investigador, sino en reorganizar estas actividades alrededor de nuevas posibilidades de procesamiento. Buscar con IA exige saber delimitar; mapear exige interpretar; sintetizar exige conservar contexto; detectar patrones exige contrastarlos; y verificar exige regresar continuamente a la fuente original. Bajo estas condiciones, la automatización puede

ampliar la capacidad de trabajar con información científica sin debilitar los principios de rigor, trazabilidad y responsabilidad que permiten convertir una colección de documentos en conocimiento científicamente defendible.

Bibliografía

- Acevedo, J. A., García, A., Aragón, M. D., y M., O. J. (2017). Modelos científicos: significado y papel en la práctica científica. *Revista científica*, 30(3), 155-166. <https://doi.org/10.14483/23448350.12288>
- Aiumtrakul, N., Thongprayoon, C., Suppadungsuk, S., Krisanapan, P., Miao, J., Qureshi, F., y Cheungpasitporn, W. (2023). Navigating the landscape of personalized medicine: The relevance of ChatGPT, BingChat, and Bard AI in nephrology literature searches. *Journal of Personalized Medicine*, 13(10). <https://doi.org/10.3390/jpm13101457>
- Ajith, A., Xia, M., Chevalier, A., Goyal, T., Chen, D., y Gao, T. (2024). LitSearch: A retrieval benchmark for scientific literature search. *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, 15068-15083. <https://doi.org/10.18653/v1/2024.emnlp-main.840>
- Arias, F., y Artigas, W. (2022). Cómo plantear problemas científicos relevantes identificando brechas de investigación. *Mujer Andina*, 1(1), 76-82. <https://doi.org/10.36881/ma.v1i1.644>
- Athaluri, S., Manthena, S., Kesapragada, V., Yarlagadda, V., Dave, T., y Duddumpudi, R. (2023). Exploring the boundaries of reality: Investigating the phenomenon of artificial intelligence hallucination in scientific writing through ChatGPT references. *Cureus*, 15(4). <https://doi.org/10.7759/cureus.37432>
- Ávila, E. (2024). Visualización de datos abiertos de investigación mediante grafos de conocimiento: metodología y aplicación de DDI-RDF y DataCite Ontology. *Scire: Representación Y organización Del Conocimiento*, 30(2), 59-71. <https://doi.org/10.54886/scire.v30i2.4963>
- Barradas, J. (2023). Inteligencia artificial como elemento transformador de la investigación científica. *Entre líneas*, 2(1), 113-122. <https://doi.org/10.56368/Entrelineas213>
- Brack, A., Hoppe, A., Stocker, M., Auer, S., y Ewerth, R. (2022). Analysing the requirements for an Open Research Knowledge Graph: Use cases, quality requirements, and construction strategies. *International Journal on Digital Libraries*, 23, 33-55. <https://doi.org/10.1007/s00799-021-00306-x>

- Castillo, W. (2023). The importance of human supervision in the use of ChatGPT as a support tool in scientific writing. *Metaverse Basic and Applied Research*, 2. <https://doi.org/10.56294/mr202329>
- Cedeño, J., Maitta, I., Vélez, M., y Palomeque, J. (2024). Investigación universitaria con inteligencia artificial. *Revista Venezolana De Gerencia*, 29(106), 817-830. <https://doi.org/10.52080/rvgluz.29.106.23>
- Chagnon, E., Pandolfi, R., Donatelli, J., y Ushizima, D. (2024). Benchmarking topic models on scientific articles using BERTeity. *Natural Language Processing Journal*, 6. <https://doi.org/10.1016/j.nlp.2023.100044>
- Chelli, M., Descamps, J., Lavoué, V., Trojani, C., Azar, M., Deckert, M., . . . Ruetsch, C. (2024). Hallucination rates and reference accuracy of ChatGPT and Bard for systematic reviews: Comparative analysis. *Journal of Medical Internet Research*, 26. <https://doi.org/10.2196/53164>
- Ciobanu, O., Aicher, A., Kernbach, J., Regli, L., Serra, C., y Staartjes, V. (2024). A critical moment in machine learning in medicine: on reproducible and interpretable learning. *Acta Neurochirurgica*, 166. <https://doi.org/10.1007/s00701-024-05892-8>
- Contreras, L., Puma, I., Morales, J., Gil, C., y Chalco, N. (2024). Ética en el uso de la inteligencia artificial en la investigación científica: desafíos y consideraciones. *Aula virtual*, 5(12), 1-22. <https://doi.org/10.5281/zenodo.14653199>
- De la Cruz, Y. (2017). El rigor científico: una necesidad de las investigaciones en ciencias de la educación. *Mikarimin. Revista Científica Multidisciplinaria*, 3(2), 41-50. <https://revista.uniandes.edu.ec/ojs/index.php/mikarimin/article/view/414>
- Dennstädt, F., Zink, J., Putora, P., Hastings, J., y Cihoric, N. (2024). Title and abstract screening for literature reviews using large language models: An exploratory study in the biomedical domain. *Systematic Reviews*, 13. <https://doi.org/10.1186/s13643-024-02575-4>
- Díaz, L. (2024). El uso de la inteligencia artificial en la investigación científica. *Revista Historia De La Educación Latinoamericana*, 26(43), 253-272. <https://doi.org/10.19053/uplc.01227238.18014>
- Doneva, S., Qin, S., Sick, B., Ellendorff, T., Goldman, J., Schneider, G., y Ineichen, B. (2024). Large language models to process, analyze, and synthesize biomedical texts: a scoping

review. *Discover Artificial Intelligence*, 4. <https://doi.org/10.1007/s44163-024-00197-2>

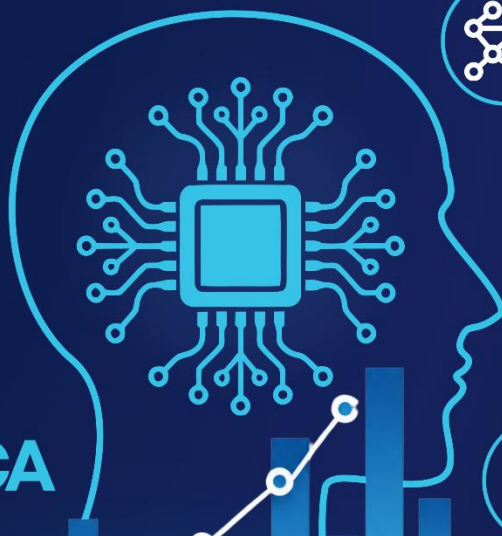
- Gallent, C., Zapata, A., y Ortego, J. (2023). El impacto de la inteligencia artificial generativa en educación superior: una mirada desde la ética y la integridad académica. *RELIEVE. Revista Electrónica de Investigación y Evaluación Educativa*, 29(2), 1-21. <https://doi.org/10.30827/relieve.v29i2.29134>
- Gao, C., Howard, F., Markov, N., Dyer, E., Ramesh, S., Luo, Y., y Pearson, A. (2023). Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. *NPJ digital medicine*, 6. <https://doi.org/10.1038/s41746-023-00819-6>
- Gilbert, S., Kather, J., y Hogan, A. (2024). Augmented non-hallucinating large language models as medical information curators. *npj Digital Medicine*, 7(100). <https://doi.org/10.1038/s41746-024-01081-0>
- Gottardi, T., Bauzer, C., y Dos, J. (2021). Semantic search to foster scientific findability: A systematic literature review. *Journal of Information and Data Management*, 12(5). <https://doi.org/10.5753/jidm.2021.1919>
- Haunschild, R., y Bornmann, L. (2024). The use of OpenAlex to produce meaningful bibliometric global overlay maps of science on the individual, institutional, and national levels. *PLOS ONE*, 19(12). <https://doi.org/10.1371/journal.pone.0308041>
- Hua, C., Cao, X., Liao, B., y Li, S. (2023). Advances on intelligent algorithms for scientific computing: An overview. *Frontiers in Neurorobotics*, 17. <https://doi.org/10.3389/fnbot.2023.1190977>
- Jaradeh, M., Singh, K., Stocker, M., Both, A., y Auer, S. (2023). Information extraction pipelines for knowledge graphs. *Knowledge and Information Systems*, 65, 1989-2016. <https://doi.org/10.1007/s10115-022-01826-x>
- Kapoor, S., y Narayanan, A. (2023). Leakage and the reproducibility crisis in machine-learning-based science. *Patterns*, 4(9). <https://doi.org/10.1016/j.patter.2023.100804>
- Krenn, M., Pollice, R., Guo, S., Aldeghi, M., Cervera, A., Friederich, P., . . . Aspuru, A. (2022). On scientific understanding with artificial intelligence. *Nature Reviews Physics*, 4, 761-769. <https://doi.org/10.1038/s42254-022-00518-3>

- Loza, J., Riera, V., Miranda, M., y Lopez, D. (2026). Reproducibilidad en los flujos de trabajo bioinformáticos aplicados a genómica y oncología clínica. *La Ciencia al Servicio de la Salud y Nutrición*, 16(2), 155-162. <https://doi.org/10.47187/cssn.Vol16.Iss2.453>
- Mata, K., Sancán, V., Káiser, I., y Kaiser, R. (2024). Una revisión sistemática del uso de la Inteligencia artificial en el desarrollo de investigaciones científicas. *Reincisol*, 3(6), 1642-1660. [https://doi.org/10.59282/reincisol.V3\(6\)1642-1660](https://doi.org/10.59282/reincisol.V3(6)1642-1660)
- Mugaanyi, J., Cai, L., Cheng, S., Lu, C., y Huang, J. (2024). Evaluation of large language model performance and reliability for citations and references in scholarly writing: Cross-disciplinary study. *Journal of Medical Internet Research*, 26. <https://doi.org/10.2196/52935>
- Mujica, R. (2021). Alfabetización Científica: Herramienta Indispensable en la Era Digital. *Revista Docentes 2.0*, 11(1), 19-26. <https://doi.org/10.37843/rtd.v11i1.181>
- Ofori, R., Aceves, M., Wiratunga, N., y Moreno, C. (2024). Towards the automation of systematic reviews using natural language processing, machine learning, and deep learning: A comprehensive review. *Artificial Intelligence Review*, 57. <https://doi.org/10.1007/s10462-024-10844-w>
- Orduña, E., y Costas, R. (2021). Link-based approach to study scientific software usage: The case of VOSviewer. *Scientometrics*, 126, 8153-8186. <https://doi.org/10.1007/s11192-021-04082-y>
- Penabad, L., Morera, M., y Penabad, M. (2024). Guía para uso y reporte de inteligencia artificial en revistas científico-académicas. *Revista Electrónica Educare*, 28(S), 1-41. <https://doi.org/10.15359/ree.28-s.19830>
- Polo, L., y Casique, R. (2025). Propuesta metodológica para la recuperación de información documental: integración de grafos de conocimiento y redes neuronales. *Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información*, 39(105), 141-163. <https://doi.org/10.22201/iibi.24488321xe.2025.105.59051>
- Ros, P., y Pérez, Á. (2023). ChatGPT: una novedosa herramienta de escritura para artículos científicos, pero no un autor (por el momento). *Revista de Neurología*, 76(8). <https://doi.org/10.33588/rn.7608.2023066>

- Ross, T., Reichmann, S., Cole, N. L., Fessler, A., Klebel, T., y Pontika, N. (2022). Dynamics of cumulative advantage and threats to equity in open science: A scoping review. *Royal Society Open Science*, 9(1), 211032. <https://doi.org/10.1098/rsos.211032>
- Semmelrock, H., Ross, T., Kopeinik, S., Theiler, D., Haberl, A., Thalmann, S., y Kowald, D. (2025). Reproducibility in machine-learning-based research: Overview, barriers and drivers. *AI Magazine*, 46(2), 1-19. <https://doi.org/10.1002/aaai.70002>
- Tamayo, O. E., Zona, R., y Loaiza, Y. E. (2015). El pensamiento crítico en la educación. Algunas categorías centrales en su estudio. *Revista Latinoamericana de Estudios Educativos*, 11(2), 111-133. <https://www.redalyc.org/pdf/1341/134146842006.pdf>
- Tang, L., Sun, Z., Idnay, B., Nestor, J., Soroush, A., Elias, P., . . . Peng, Y. (2023). Evaluating large language models on medical evidence summarization. *npj Digital Medicine*, 6. <https://doi.org/10.1038/s41746-023-00896-7>
- Tapia, R., y Recalde, A. (2024). Mejora de gestión documental: revisión sistemática. Ingenium et Potentia. *Revista Electrónica Multidisciplinaria de Ciencias Básicas, Ingeniería y Arquitectura*, 6(11), 16-26. <https://doi.org/10.35381/i.p.v6i11.4076>
- Torre, J. d., Ramírez, A., y Romero, J. (2023). Artificial intelligence to automate the systematic review of scientific literature. *Computing*, 105, 2171-2194. <https://doi.org/10.1007/s00607-023-01181-x>
- Vázquez, M., Pereira, J., Cid, J., y Arenas, J. (2022). Validation of scientific topic models using graph analysis and corpus metadata. *Scientometrics*, 127, 5441-5458. <https://doi.org/10.1007/s11192-022-04318-5>
- Verma, S., Bhatia, R., Harit, S., y Batish, S. (2023). Scholarly knowledge graphs through structuring scholarly communication: A review. *Complex & Intelligent Systems*, 9, 1059-1095. <https://doi.org/10.1007/s40747-022-00806-6>
- Wadden, D., Lo, K., Kuehl, B., Cohan, A., Beltagy, I., Wang, L., y Hajishirzi, H. (2022). SciFact-Open: Towards open-domain scientific claim verification. *Findings of the Association for Computational Linguistics: EMNLP 2022*, 4719-4734. <https://doi.org/10.18653/v1/2022.findings-emnlp.347>
- Wadden, D., y Lo, K. (2021). Overview and insights from the SCIVER shared task on scientific claim verification. *Proceedings of the Second Workshop on Scholarly Document Processing*, 124-129. <https://aclanthology.org/2021.sdp-1.16.pdf>

- Walters, W., y Wilder, E. (2023). Fabrication and errors in the bibliographic citations generated by ChatGPT. *Scientific Reports*, 13. <https://doi.org/10.1038/s41598-023-41032-5>
- Wu, X., Nguyen, T., y Luu, A. (2024). A survey on neural topic models: Methods, applications, and challenges. *Artificial Intelligence Review*, 57. <https://doi.org/10.1007/s10462-023-10661-7>
- Zárate, L. O. (2022). Explicabilidad (de la inteligencia artificial). *EUNOMÍA. Revista en Cultura de la Legalidad*(22), 328-344. <https://doi.org/10.20318/eunomia.2022.6819>
- Zhang, Z., Patra, B., Yaseen, A., Zhu, J., Sabharwal, R., Roberts, K., . . . Wu, H. (2023). Scholarly recommendation systems: A literature survey. *Knowledge and Information Systems*, 65, 4433-4478. <https://doi.org/10.1007/s10115-023-01901-x>

NUEVAS FORMAS DE
BUSCAR
Y ANALIZAR
CON IA LA
INFORMACIÓN CIENTÍFICA



CIDPROS
EDITORIAL

ISBN: 978-9907-9574-8-8



9 789907 957488