

INVESTIGACIÓN CIENTÍFICA en la era de la INTELIGENCIA ARTIFICIAL

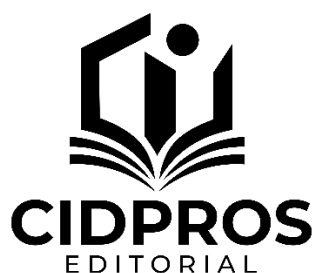


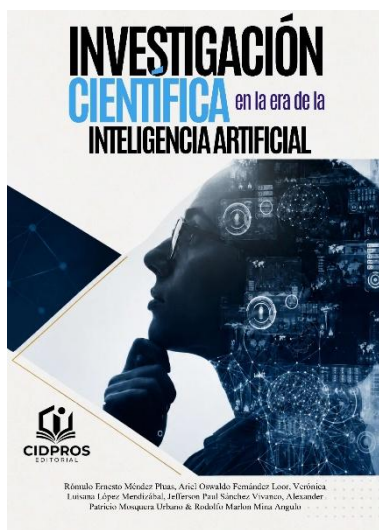

CIDPROS
EDITORIAL

Rómulo Ernesto Méndez Pluas, Ariel Oswaldo Fernández Loor, Verónica
Luisana López Mendizábal, Jefferson Paul Sánchez Vivanco, Alexander
Patricio Mosquera Urbano & Rodolfo Marlon Mina Angulo

Investigación científica en la era de la inteligencia artificial

Rómulo Ernesto Méndez Pluas, Ariel Oswaldo
Fernández Loor, Verónica Luisana López
Mendizábal, Jefferson Paul Sánchez Vivanco,
Alexander Patricio Mosquera Urbano & Rodolfo
Marlon Mina Angulo





Datos bibliográficos

ISBN	978-9907-848-06-9
Título del libro	Investigación Científica en la Era de la Inteligencia Artificial
Autores	Rómulo Ernesto Méndez Pluas Ariel Oswaldo Fernández Loo Verónica Luisana López Mendizábal Jefferson Paul Sánchez Vivanco Alexander Patricio Mosquera Urbano Rodolfo Marlon Mina Angulo
Editorial	CIDPROS EDITORIAL
Materia	001.4 - Investigación
Público objetivo	Profesional / académico
Año	2026
Número de edición	1
Tamaño	3.46Mb
Soporte	Libro digital descargable
Formato	PDF (.pdf)
Idioma	Español
DOI	https://doi.org/10.70577/pd8w9r97
	Hecho en Ecuador / Made in Ecuador

Esp. Rómulo Ernesto Méndez Pluas

Investigador Independiente
rmendez_pluas@yahoo.com
<https://orcid.org/0009-0005-2745-8251>
Santa Elena, Santa Elena, Ecuador.

Semblanza

Rómulo Ernesto Méndez Pluas es licenciado en enfermería y prácticas naturistas, ecuatoriano nacido el 11 de noviembre de 1983 en Santa Elena, residente en la ciudad de Quito desde hace una década, donde nace la vocación del cuidado a los pacientes en estado muy crítico de salud.

Cuenta con una sólida formación académica, realizó el Diplomado en Manejo Avanzado de Heridas y Ostomías, cuenta con la Certificación Multifiltrate Terapia CRRT Hipoperfusión y Anticoagulación CICA, certificación en Soporte Vital Cardiovascular Avanzado – ACLS, es Licenciado en Enfermería y Etnoterapias, Especialista en Enfermería en Medicina Crítica y Magister y posterior obtuvo el título Magister de Educación con mención en Docencia e Investigación en Educación Superior. Esta preparación refleja su vocación por la excelencia académica y la actualización permanente en el área de la salud y la educación superior.



Ha dedicado años a la atención del paciente en los servicios de emergencias, cardiología, nefrología y terapia intensiva en los principales hospitales de la capital y el país, también hasta la actualidad, ha dedicado un espacio para ser docente y tutor de prácticas hospitalarias para estudiantes de pre y posgrados de la carrera de enfermería de las principales universidades del país.

MSc. Ariel Oswaldo Fernández Loor

Universidad Estatal Península de Santa Elena

arialdo999@gmail.com

<https://orcid.org/0000-0003-0508-6133>

Quevedo, Los Ríos, Ecuador

Semblanza

Ariel Oswaldo Fernández Loor es un ingeniero en sistemas e investigador ecuatoriano nacido el 28 de febrero de 2000. Su trayectoria profesional se ha desarrollado entre los cantones Portoviejo, Quito y Santa Elena, consolidando una visión enfocada en la innovación tecnológica, el desarrollo de software y el uso de la ciencia de datos como motores clave para la transformación digital e institucional.

Cuenta con una sólida formación académica: es Ingeniero en Sistemas por la Universidad Técnica Estatal de Quevedo y Magíster en Telecomunicaciones por la Universidad Estatal Península de Santa Elena. Ha complementado su perfil con capacitaciones especializadas en Inteligencia Artificial aplicada a la Educación, Cultura de la Investigación y redes avanzadas, reflejando un firme compromiso con la excelencia científica y la actualización profesional continua.

En el ámbito laboral, se ha desempeñado como analista de desarrollo de sistemas, técnico de plataformas virtuales y docente técnico, además de haber sido ayudante de cátedra universitaria. Como investigador, ha publicado artículos en revistas de prestigio internacional (Scopus/SciELO) en las áreas de IoT, redes inteligentes y analítica de datos, orientando su labor hacia la docencia y el desarrollo de soluciones de alto impacto social.



Mgt. Verónica Luisana López Mendizábal

Instituto Superior Tecnológico Babahoyo
vlopez@istb.edu.ec
<https://orcid.org/0009-0003-9444-5843>
Babahoyo, Los Rios, Ecuador

Semblanza

Verónica Luisana López Mendizábal nació el 11 de marzo de 1983 en el cantón Babahoyo, provincia de Los Ríos, Ecuador, ciudad donde creció y consolidó los valores que han guiado su vida personal y profesional. Es Ingeniera en Sistemas por la Universidad Técnica de Babahoyo (2009). Posteriormente obtuvo el título de Magíster en Docencia y Currículo por la Universidad Técnica de Babahoyo (2016), fortaleciendo sus competencias en pedagogía, planificación educativa y gestión del aprendizaje. En el año 2023 alcanzó el título de Magíster en Sistemas de Información, mención en Inteligencia de Negocios y Analítica de Datos Masivos, otorgado por la Universidad Estatal de Milagro, consolidando una formación multidisciplinaria que integra la tecnología, la analítica de datos y la gestión estratégica para la toma de decisiones institucionales.



Desde el año 2015 se desempeña como docente del Instituto Superior Tecnológico Babahoyo, institución en la que ha contribuido significativamente a la formación de profesionales en el área de Desarrollo de Software. Durante su trayectoria académica ha impartido asignaturas como Fundamentos de Programación, Base de Datos, Estadística Descriptiva, Análisis y Diseño de Sistemas, Ofimática, Proyectos de Investigación y otras relacionadas con las tecnologías de la información, promoviendo metodologías activas de aprendizaje, la innovación educativa y el fortalecimiento de las competencias profesionales de sus estudiantes.

Su experiencia trasciende el ámbito docente, destacándose en la gestión académica, administrativa y estratégica de la educación superior. Ha liderado importantes procesos institucionales como Responsable de la Unidad de Planificación Estratégica, impulsando la formulación, actualización, seguimiento y evaluación del Plan Estratégico de Desarrollo Institucional (PEDI) y del Plan Operativo Anual (POA), así como el diseño de instrumentos de planificación, control y mejora continua. Asimismo, ha participado activamente en procesos de aseguramiento de la calidad y acreditación institucional.

MSc. Jefferson Paul Sánchez Vivanco

Universidad Estatal Península de Santa Elena (UPSE)

jeffer2226@gmail.com

<https://orcid.org/0009-0001-8566-795X>

Santa Elena, Santa Elena, Ecuador

Semblanza

Jefferson Paul Sánchez Vivanco es un ingeniero en sistemas informáticos ecuatoriano, nacido el 12 de febrero de 1995 y residente en Santa Elena, Ecuador. Está vinculado a la Universidad Estatal Península de Santa Elena (UPSE), donde orienta su actividad profesional al desarrollo y mejora de soluciones tecnológicas para la gestión institucional. Su perfil combina la ingeniería de software, la administración de datos y la aplicación responsable de tecnologías digitales en contextos académicos y administrativos.



Es Ingeniero en Sistemas Informáticos por la Universidad Técnica de Manabí. Posteriormente obtuvo el grado de Magíster en Computación, mención en Sistemas Inteligentes, en la Escuela Politécnica Nacional, y el título de Máster Universitario en Ciberseguridad por la Universidad Internacional de La Rioja. Esta formación le ha permitido fortalecer competencias en desarrollo de software, bases de datos, inteligencia aplicada, análisis de riesgos y protección de sistemas de información.

A lo largo de su trayectoria ha participado en el diseño, automatización y optimización de sistemas orientados a procesos universitarios, integrando buenas prácticas de seguridad, análisis de información y mejora continua. Sus intereses investigativos se concentran en la ciberseguridad, el Internet de las cosas y el estudio de cámaras IP, con énfasis en la identificación de vulnerabilidades y la formulación de propuestas de protección. Se caracteriza por su compromiso con el aprendizaje permanente, la ética profesional y el uso de la tecnología para generar soluciones seguras, eficientes y útiles para la sociedad.

PhD. Alexander Patricio Mosquera Urbano

Universidad central del Ecuador - Figempa

apmosquerau@uce.edu.ec

<https://orcid.org/0009-0003-9860-213X>

Quito, Pichincha, Ecuador

Semblanza

Alexander Patricio Mosquera Urbano es un académico y profesional ecuatoriano, nacido el 4 de julio de 1991 en la ciudad de Otavalo, provincia de Imbabura. Desde temprana edad ha mantenido un profundo vínculo con su tierra y un marcado compromiso con la sociedad. Su vocación de liderazgo le permitió representar internacionalmente al Ecuador en los Estados Unidos de América, fue becario de la Federación de Rusia, Gobiernos de Bélgica, su experiencia contribuyó a fortalecer su visión intercultural, su compromiso con el desarrollo del país y, posteriormente, su vocación por la docencia universitaria y la formación de nuevas generaciones.



En el ámbito académico, posee un PhD. en Ciencias Económicas, un Máster en Administración y el título de Ingeniero Eléctrico (Bachelor) otorgado por la Universidad de Minería de San Petersburgo Emperatriz Catalina II, de la Federación de Rusia. Como parte de su compromiso permanente con la actualización profesional y los nuevos desafíos tecnológicos, actualmente cursa una segunda maestría en Inteligencia Artificial y Ciencia de Datos, fortaleciendo un perfil multidisciplinario que integra ingeniería, economía, administración y tecnologías emergentes.

Su trayectoria profesional se ha desarrollado principalmente en los ámbitos de la docencia universitaria, investigación científica, energías renovables, minería e innovación tecnológica. Actualmente se desempeña como profesor universitario en la Universidad Central del Ecuador y como director académico del Instituto de Innovación Tecnológica, Energía y Minas. Asimismo, ejerce actividades docentes en diferentes instituciones de educación superior a nivel nacional y participa activamente en iniciativas académicas, científicas y de investigación.

MSc. Rodolfo Marlon Mina Angulo

Universidad Técnica Luis Vargas Torres
marlon.mina.angulo@utelvt.edu.ec
<https://orcid.org/0000-0002-5823-1034>
Esmeraldas, Esmeraldas, Ecuador.

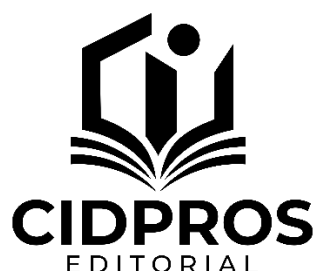
Semblanza

Rodolfo Marlon Mina Angulo es un docente nacido el 25 de agosto de 1976 en Esmeraldas, cuya vida ha estado marcada por su residencia en el cantón Esmeraldas, un entorno urbano que fortaleció su amor por la educación y su visión humanista de la docencia como instrumento al servicio del bien común. Desde temprana edad desarrolló una profunda empatía por la situación de la juventud y adolescencia lo que ha orientado su compromiso con este sector de la sociedad esmeraldeña y ecuatoriana, la equidad y el desarrollo social.



Cuenta con una sólida formación académica: es Tecnólogo Programador de Sistemas, Licenciado en Informática Educativa y catedrático de la Universidad Técnica Luis Vargas Torres de Esmeraldas desde 2018. Posteriormente obtuvo el título de Magíster en Ingeniería de Software y Sistemas Informáticos en España. Esta preparación refleja su vocación por la excelencia académica y la actualización permanente en las ciencias pedagógicas.

A lo largo de su trayectoria, ha destacado tanto en su vida personal como en el servicio público, promoviendo principios, valores, equidad y con enfoque de vida, basada en la dignidad humana y el respeto a los derechos fundamentales de una sociedad más justa y equitativa. Su vocación académica se evidencia en artículos como “Estrategias Innovadoras Para Mitigar la Suplantación de Identidad en Redes Sociales” y el libro “Principios Fundamentales para la Iniciación de la Investigación Científica.”. Padre y profesional comprometido, representa valores de resiliencia, ética e integridad, inspirando a nuevas generaciones a construir un Ecuador más justo y solidario donde la juventud sean más críticos, autocríticos, reflexiva y analítica.



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).

Derechos de Autor © 2026

Este documento se publica bajo los términos y condiciones de la licencia:



Creative Commons Reconocimiento – No Comercial – Compartir Igual 4.0 Internacional (CC BY-NC-SA 4.0)

Usted es libre de compartir y adaptar el material, siempre que otorgue el crédito correspondiente, no lo utilice con fines comerciales y distribuya sus contribuciones bajo la misma licencia.

Para ver una copia de esta licencia, visite: <https://creativecommons.org/licenses/by-nc-sa/4.0/deed.es>

Sinopsis

Investigación Científica en la Era de la Inteligencia Artificial analiza la transformación de los procesos de producción del conocimiento ante la incorporación de modelos generativos, aprendizaje automático, sistemas multiagente, laboratorios autónomos y herramientas de automatización científica. La obra examina cómo estas tecnologías intervienen en la exploración de literatura, formulación de hipótesis, análisis de datos, programación, experimentación, redacción y evaluación de resultados.

A lo largo de cinco capítulos se estudia la evolución desde la computación científica hasta la inteligencia artificial para la ciencia; la reconfiguración del método científico; la colaboración humano-máquina; el descubrimiento asistido por modelos computacionales; y las competencias necesarias para investigar en entornos inteligentes. También se desarrollan la automatización de revisiones sistemáticas, la preparación de datos, los análisis cualitativos y multimodales, los agentes autónomos y la ciencia de ciclo cerrado.

Desde una perspectiva crítica, el libro aborda validez, confiabilidad, reproducibilidad, explicabilidad, sesgos algorítmicos, alucinaciones, integridad científica, autoría y revisión por pares. Asimismo, examina la ciencia abierta, la inferencia causal, los ecosistemas de datos reutilizables y la gobernanza responsable de sistemas con distintos niveles de autonomía.

La obra sostiene que la inteligencia artificial puede ampliar la escala, velocidad y capacidad analítica de la investigación, pero no sustituye la interpretación disciplinar, la comprobación empírica ni la responsabilidad humana. Su incorporación científica exige trazabilidad, validación independiente, transparencia y mecanismos institucionales capaces de garantizar un uso riguroso, ético y verificable.

Palabras clave: inteligencia artificial; investigación científica; automatización; integridad científica; gobernanza.

Synopsis

Scientific Research in the Age of Artificial Intelligence analyzes the transformation of knowledge-production processes resulting from the incorporation of generative models, machine learning, multi-agent systems, autonomous laboratories, and scientific automation tools. The book examines how these technologies intervene in literature exploration, hypothesis formulation, data analysis, programming, experimentation, writing, and result evaluation.

Across five chapters, it studies the evolution from scientific computing to artificial intelligence for science; the reconfiguration of the scientific method; human-machine collaboration; discovery assisted by computational models; and the competencies required to conduct research in intelligent environments. It also addresses the automation of systematic reviews, data preparation, qualitative and multimodal analysis, autonomous agents, and closed-loop science.

From a critical perspective, the book examines validity, reliability, reproducibility, explainability, algorithmic bias, hallucinations, scientific integrity, authorship, and peer review. It also explores open science, causal inference, reusable data ecosystems, and the responsible governance of systems with different levels of autonomy.

The book argues that artificial intelligence can expand the scale, speed, and analytical capacity of research, but it does not replace disciplinary interpretation, empirical verification, or human responsibility. Its scientific incorporation requires traceability, independent validation, transparency, and institutional mechanisms capable of ensuring rigorous, ethical, and verifiable use.

Keywords: artificial intelligence; scientific research; automation; scientific integrity; governance.

Contenido

Sinopsis.....	11
Synopsis.....	12
Introducción.....	15
Capítulo 1 Transformación de la investigación científica mediante inteligencia artificial	18
1.1 De la computación científica a la inteligencia artificial para la ciencia.....	19
1.2 Reconfiguración del método científico	22
1.3 Inteligencia aumentada y colaboración humano-máquina.....	26
1.4 Descubrimiento científico asistido por modelos computacionales	29
1.5 Transformaciones disciplinares y convergencia científica	33
1.6 Infraestructuras digitales para la investigación inteligente	37
1.7 Competencias del investigador en entornos asistidos por inteligencia artificial ..	41
Capítulo 2 Generación del conocimiento con inteligencia artificial	45
2.1 Exploración automatizada de literatura científica.....	46
2.2 Generación y priorización de hipótesis	49
2.3 Modelos generativos como instrumentos de razonamiento científico	53
2.4 Descubrimiento de patrones en datos complejos	56
2.5 Integración de conocimiento basado en datos y teoría.....	60
2.6 Simulación, predicción y construcción de escenarios.....	64
2.7 Producción de conocimiento en investigación cualitativa	68
Capítulo 3 Automatización de los procesos de investigación científica	73
3.1 Automatización de revisiones sistemáticas.....	74
3.2 Recolección y preparación automatizada de datos	77
3.3 Análisis estadístico y programación asistida	81
3.4 Laboratorios autónomos y experimentación robótica	84
3.5 Automatización de análisis cualitativo y multimodal	89
3.6 Redacción, visualización y comunicación asistidas.....	93
3.7 Agentes autónomos y automatización integral del ciclo científico.....	97
Capítulo 4 Calidad y transparencia en investigaciones con inteligencia artificial	102
4.1 Validez y confiabilidad de resultados asistidos por inteligencia artificial.....	103
4.2 Reproducibilidad y trazabilidad computacional.....	107
4.3 Explicabilidad e interpretabilidad de los modelos	111
4.4 Sesgos algorítmicos y representatividad de los datos	114

4.5 Alucinaciones, errores y verificación de resultados generativos	119
4.6 Integridad científica, autoría y declaración del uso de inteligencia artificial	123
4.7 Revisión por pares y evaluación científica asistidas por inteligencia artificial ..	127
Capítulo 5 Nuevos paradigmas de investigación en la era digital	132
5.1 Ciencia abierta y ecosistemas de datos reutilizables	133
5.2 Investigación basada en agentes y sistemas multiagente	136
5.3 Laboratorios autónomos y ciencia de ciclo cerrado	140
5.4 Investigación multimodal y conocimiento integrado	144
5.5 Inferencia causal y descubrimiento de mecanismos	147
5.6 Gobernanza responsable de la inteligencia artificial científica	152
5.7 Perspectivas futuras de la investigación científica	156
Capítulo 6 Implicaciones epistemológicas, profesionales y sociales de la investigación con inteligencia artificial	161
6.1 Autoridad epistémica y juicio científico en entornos asistidos por inteligencia artificial	162
6.2 Autonomía intelectual y dependencia cognitiva del investigador	165
6.3 Mediación algorítmica en la selección de problemas y agendas científicas	168
6.4 Homogeneización del conocimiento y preservación de la diversidad científica	172
6.5 Asimetrías de acceso y concentración de capacidades científicas	176
6.6 Transformación del trabajo investigador y de la experticia científica	179
6.7 Confianza y legitimidad del conocimiento científico mediado por inteligencia artificial	183
Bibliografía	188

Introducción

La inteligencia artificial se ha incorporado progresivamente a la investigación científica como una tecnología capaz de procesar información, reconocer patrones, generar contenidos y coordinar tareas complejas. Su presencia ya no se limita a operaciones estadísticas o simulaciones previamente programadas, sino que alcanza la búsqueda documental, la formulación de hipótesis, la preparación de datos, la experimentación y la comunicación de resultados. Esta ampliación funcional modifica la manera en que se organizan los proyectos científicos y plantea nuevas exigencias para la validación, la interpretación y la responsabilidad sobre el conocimiento producido.

La transformación actual debe comprenderse como parte de una evolución más amplia de la computación científica. Los sistemas informáticos comenzaron apoyando cálculos, modelizaciones y simulaciones, para posteriormente incorporar reglas expertas, aprendizaje basado en datos y modelos generativos capaces de responder a instrucciones en lenguaje natural. En la actualidad, ciertas arquitecturas pueden conectarse con repositorios, programas, bases de datos e instrumentos experimentales. Este tránsito no convierte automáticamente a las máquinas en investigadoras, pero sí amplía su participación dentro de procesos anteriormente desarrollados casi por completo por personas.

La incorporación de estas herramientas reconfigura el método científico sin eliminar sus fundamentos. La observación, la formulación de problemas, la revisión de antecedentes, la generación de hipótesis, el diseño metodológico, el análisis y la comunicación continúan requiriendo coherencia, contrastación y revisión crítica. La inteligencia artificial puede ampliar la escala y velocidad de estas actividades, pero sus resultados dependen de los datos disponibles, los criterios incorporados y las instrucciones utilizadas. Por ello, toda salida algorítmica debe considerarse provisional hasta que sea examinada mediante procedimientos metodológicos y evidencia independiente.

Esta obra adopta una perspectiva de inteligencia aumentada, entendida como una relación de complementariedad entre capacidades humanas y computacionales. Los sistemas automatizados aportan velocidad, consistencia operativa y procesamiento a gran escala, mientras los investigadores proporcionan conocimiento disciplinar, creatividad, interpretación contextual y juicio crítico. La colaboración resulta productiva cuando existe una distribución clara de funciones y cuando las decisiones científicas sustantivas

permanecen bajo control humano. Delegar tareas no significa transferir la responsabilidad por los errores, las conclusiones o las consecuencias derivadas de su utilización.

El libro también reconoce que la inteligencia artificial no transforma de igual manera todas las disciplinas. En salud puede integrar datos clínicos, imágenes y señales; en química y materiales puede apoyar el diseño y la experimentación; en ingeniería permite simular y optimizar procesos; y en educación y ciencias sociales amplía el análisis de trayectorias, discursos y comportamientos. Sin embargo, cada campo mantiene criterios propios de evidencia y validación. La utilidad de una herramienta depende de su adecuación al fenómeno estudiado y no únicamente de su sofisticación técnica.

Uno de los propósitos centrales de la obra es diferenciar la capacidad para generar resultados de la capacidad para producir conocimiento científicamente válido. Un modelo puede identificar asociaciones, elaborar predicciones, redactar explicaciones o proponer hipótesis sin demostrar que sus conclusiones sean verdaderas. La precisión tampoco debe confundirse con validez, ni la coherencia lingüística con comprensión científica. Para evaluar adecuadamente las salidas asistidas por inteligencia artificial se requieren controles de robustez, reproducibilidad, trazabilidad, generalización y correspondencia con teorías, datos y procedimientos de comprobación.

La automatización creciente introduce, además, riesgos que deben ser tratados como problemas metodológicos y no solo como advertencias éticas generales. Las referencias inexistentes, los datos fabricados, los errores de código, los sesgos de selección y la pérdida de contexto pueden propagarse entre etapas conectadas. Del mismo modo, el uso de información confidencial, la falta de declaración de herramientas y la dependencia de plataformas externas comprometen la integridad de los proyectos. Frente a estas dificultades, la obra propone verificación con fuentes primarias, documentación de versiones, auditoría y supervisión proporcional al riesgo.

La organización del libro responde a esta visión integral. El primer capítulo examina la transformación de la investigación y las competencias requeridas; el segundo analiza la generación del conocimiento mediante exploración documental, hipótesis, patrones, simulación e investigación cualitativa; el tercero estudia la automatización de procedimientos; el cuarto desarrolla criterios de calidad, transparencia e integridad; y el quinto aborda ciencia abierta, sistemas multiagente, laboratorios de ciclo cerrado,

multimodalidad, inferencia causal, gobernanza y escenarios futuros. Esta secuencia conecta posibilidades técnicas con condiciones metodológicas e institucionales.

La obra está dirigida a investigadores, docentes, estudiantes de posgrado, editores y profesionales interesados en comprender cómo utilizar inteligencia artificial sin debilitar los fundamentos de la actividad científica. No pretende presentar estas tecnologías como soluciones universales ni anticipar una sustitución inevitable del investigador. Su propósito es ofrecer criterios para identificar usos pertinentes, delimitar niveles de autonomía y establecer mecanismos de control. En este sentido, la innovación adquiere valor cuando contribuye a producir conocimiento verificable, reproducible, explicable y socialmente responsable.

El futuro de la investigación científica dependerá menos de la cantidad de tareas automatizadas que de la calidad de las relaciones construidas entre datos, modelos, instituciones y personas. La inteligencia artificial puede ampliar las preguntas investigables y acelerar procedimientos, pero también concentrar capacidades, homogeneizar enfoques y ocultar incertidumbres. Por ello, su integración exige competencias críticas, infraestructuras abiertas, políticas adaptables y responsabilidad humana identificable. El desafío no consiste en elegir entre ciencia humana o ciencia automatizada, sino en construir formas de colaboración que preserven el rigor, la diversidad y la confianza científica.

Capítulo

1

Transformación de la investigación científica mediante inteligencia artificial

Rómulo Ernesto Méndez Pluas



1.1 De la computación científica a la inteligencia artificial para la ciencia

La computación científica surgió como una extensión de las capacidades humanas para resolver cálculos, representar fenómenos y ejecutar simulaciones imposibles de abordar manualmente. Durante sus primeras etapas, los sistemas informáticos dependían de instrucciones explícitas, modelos matemáticos previamente definidos y procedimientos programados por especialistas. Su valor residía en aumentar la velocidad y precisión del procesamiento, sin intervenir en la formulación del conocimiento. Esta base permitió consolidar métodos numéricos, análisis estadísticos y entornos de simulación que posteriormente facilitaron la incorporación de técnicas de inteligencia artificial en la actividad científica.

El tránsito hacia la inteligencia artificial comenzó cuando los sistemas dejaron de limitarse a ejecutar secuencias rígidas y adquirieron capacidades para representar conocimiento, aplicar reglas y resolver problemas delimitados. Los sistemas expertos constituyeron una etapa relevante porque intentaron reproducir decisiones especializadas mediante bases de conocimiento y mecanismos de inferencia. Aunque dependían de reglas construidas manualmente y tenían dificultades para adaptarse a situaciones nuevas, demostraron que la computación podía apoyar tareas interpretativas y decisorias, no solo operaciones matemáticas, ampliando su función dentro de diferentes campos científicos.

La expansión del aprendizaje automático modificó este enfoque al permitir que los modelos identificaran regularidades directamente a partir de los datos. En lugar de programar cada decisión, los investigadores comenzaron a entrenar algoritmos para clasificar, predecir, reconocer imágenes, procesar lenguaje y detectar relaciones complejas. Incio et al. (2022) destacan que las redes neuronales, los macrodatos, la visión computacional y el análisis predictivo ampliaron las aplicaciones de la IA. Este cambio trasladó parte del esfuerzo científico desde la formulación de reglas hacia la selección, preparación y evaluación de datos.

El desarrollo del aprendizaje profundo fortaleció esa transformación mediante redes capaces de construir representaciones jerárquicas a partir de grandes conjuntos de información. Estas técnicas incrementaron el rendimiento en reconocimiento de voz, clasificación de imágenes y toma de decisiones complejas, áreas cuyo crecimiento reciente evidencia la evolución funcional del concepto de inteligencia artificial (Labañino et al., 2025). Sin embargo, el avance no consistió únicamente en modelos más precisos:

también cambió la escala de los problemas investigables y la posibilidad de analizar información heterogénea mediante procedimientos parcialmente automatizados.

Una nueva etapa se consolidó con los modelos generativos y fundacionales, entrenados sobre grandes volúmenes de texto, imágenes, código u otras modalidades. Estos sistemas pueden adaptarse a múltiples tareas, producir contenidos, sintetizar literatura y apoyar procesos de razonamiento mediante instrucciones en lenguaje natural. Zhang et al. (2025) explican que los grandes modelos de lenguaje pueden intervenir en distintas fases del método científico, desde la exploración de antecedentes hasta la formulación y evaluación de hipótesis. Su versatilidad transforma la interacción entre investigadores, datos y herramientas computacionales.

Pese a estas capacidades, conviene diferenciar la inteligencia artificial aplicada a la investigación de la inteligencia artificial para la ciencia. La primera utiliza herramientas generales para apoyar tareas como redactar, clasificar documentos o analizar datos. La segunda diseña sistemas orientados a problemas científicos, integrando conocimiento disciplinar, modelos computacionales, evidencia y mecanismos de validación. Krenn et al. (2022) sostienen que la contribución de la IA debe valorarse también por su capacidad para favorecer comprensión científica, no solamente por generar predicciones exactas o acelerar procedimientos previamente establecidos.

Esta distinción permite reconocer un cambio desde herramientas pasivas hacia sistemas con participación más activa en la investigación. Suazo (2023) identifica aplicaciones relacionadas con el análisis de grandes cantidades de datos, la generación y validación de hipótesis, las simulaciones y la optimización de procesos. En este contexto, la IA puede proponer relaciones, organizar alternativas y orientar decisiones preliminares. No obstante, los resultados computacionales mantienen carácter provisional hasta que son contrastados con conocimientos disciplinares, procedimientos metodológicos rigurosos y evidencia científica obtenida mediante observación o experimentación.

La evolución reciente también muestra una transición hacia sistemas conectados con bases documentales, programas, instrumentos y entornos de experimentación. Los modelos fundacionales pueden actuar como interfaces capaces de coordinar herramientas especializadas y organizar secuencias de tareas científicas. Esta capacidad anticipa el desarrollo de agentes orientados a objetivos, aunque su funcionamiento, autonomía y riesgos requieren un tratamiento específico posterior. En esta etapa introductoria, lo

relevante es comprender que la inteligencia artificial ya no opera únicamente sobre resultados finales, sino que comienza a integrarse en diferentes momentos del proceso investigador.

La evolución desde la computación científica hasta la inteligencia artificial para la ciencia representa una ampliación progresiva de funciones: calcular, simular, inferir, aprender, generar y coordinar. Cada etapa conserva aportes de las anteriores, pero introduce nuevas posibilidades y exigencias metodológicas. La capacidad de producir respuestas sofisticadas no convierte automáticamente a un sistema en generador de conocimiento válido. El valor científico de estas tecnologías depende de la calidad de los datos, la adecuación de los modelos, la interpretación experta y la comprobación de sus resultados.

Figura 1

Evolución de la computación científica hacia la inteligencia artificial para la ciencia



Nota. La figura representa la evolución funcional de los sistemas computacionales, desde el procesamiento de cálculos y simulaciones hasta su integración en tareas de generación, análisis y descubrimiento científico. Elaboración propia.

La evolución de la computación científica hacia la inteligencia artificial para la ciencia refleja una ampliación progresiva de las capacidades tecnológicas aplicadas a la investigación. Los sistemas pasaron de ejecutar cálculos y reglas predefinidas a aprender

de los datos, generar contenidos y coordinar herramientas orientadas al descubrimiento. No obstante, esta transformación no elimina el papel del investigador, sino que exige mayor capacidad para seleccionar modelos, interpretar resultados y comprobar su validez. La inteligencia artificial fortalece la producción científica cuando se integra con conocimiento disciplinar, evidencia empírica y supervisión humana responsable.

1.2 Reconfiguración del método científico

La incorporación de inteligencia artificial está reconfigurando el método científico al modificar la forma en que se observan fenómenos, se delimitan problemas y se organizan evidencias. Su influencia no implica abandonar la secuencia lógica de la investigación, sino redistribuir tareas entre capacidades humanas y sistemas computacionales. Díaz (2024) señala que estas tecnologías amplían las posibilidades para localizar información, reconocer relaciones y orientar decisiones metodológicas. El cambio principal reside en aumentar la escala y velocidad del trabajo sin eliminar la necesidad de juicio crítico y responsabilidad sobre los resultados obtenidos.

En la observación científica, la inteligencia artificial permite procesar registros extensos, imágenes, señales y datos producidos continuamente por instrumentos digitales. Esta capacidad amplía aquello que puede considerarse observable, porque facilita detectar regularidades que resultan difíciles de identificar mediante inspección directa. Sin embargo, la observación asistida depende de criterios previos sobre qué datos recoger, cómo representarlos y cuáles variables conservar. Por ello, los patrones identificados no constituyen hallazgos autosuficientes, sino resultados condicionados por decisiones de diseño que deben mantenerse explícitas, justificadas y sujetas a revisión durante todo el estudio.

La formulación del problema también adquiere una dinámica distinta cuando los investigadores utilizan sistemas capaces de sintetizar antecedentes, comparar enfoques y señalar vacíos aparentes. Ruiz (2024) sostiene que la inteligencia artificial puede intervenir desde la generación de hipótesis hasta la interpretación de resultados, favoreciendo una exploración más amplia del campo. No obstante, delimitar un problema exige valorar su relevancia científica, social y disciplinar. Esa decisión no puede derivarse únicamente de tendencias documentales, frecuencia de términos o disponibilidad de datos procesables, sino de una evaluación razonada del contexto investigado.

Durante la revisión de antecedentes, los sistemas de inteligencia artificial pueden ordenar grandes volúmenes de publicaciones y establecer relaciones temáticas con rapidez. Este apoyo facilita reconocer debates, conceptos recurrentes y áreas insuficientemente estudiadas. Sin embargo, la recuperación automatizada puede omitir investigaciones, privilegiar determinados idiomas o reproducir desigualdades presentes en las bases consultadas. En consecuencia, la revisión conserva su función crítica: evaluar la calidad de las fuentes, interpretar diferencias entre estudios y comprobar que las referencias recuperadas existen, son pertinentes y respaldan realmente las afirmaciones formuladas.

La generación de hipótesis se transforma cuando los modelos combinan literatura, datos y relaciones previas para proponer explicaciones posibles. Zhang et al. (2025) describen la participación de los grandes modelos de lenguaje en distintas fases del método científico, incluida la formulación de hipótesis. Su utilidad consiste en ampliar el conjunto de alternativas que el investigador puede considerar, no en determinar cuál es verdadera. La plausibilidad teórica, la posibilidad de contrastación y la coherencia con el conocimiento disponible continúan siendo criterios científicos bajo responsabilidad humana, mediante contraste empírico y revisión especializada.

En el diseño metodológico, la inteligencia artificial puede ayudar a comparar estrategias, anticipar dificultades y organizar procedimientos de acuerdo con los objetivos establecidos. Barradas (2023) destaca su capacidad para optimizar procesos científicos y apoyar el análisis de problemas complejos. Aun así, elegir un enfoque, una muestra, una técnica o un criterio de medición requiere comprender la naturaleza del fenómeno estudiado. La recomendación automática solo es válida cuando se contrasta con fundamentos epistemológicos, condiciones contextuales y exigencias éticas propias de cada investigación, así como con los recursos realmente disponibles.

La experimentación asistida introduce posibilidades para simular condiciones, seleccionar combinaciones prometedoras y ajustar procedimientos según resultados preliminares. Esta capacidad reduce tiempos y recursos, especialmente cuando el espacio de alternativas es demasiado amplio para una exploración manual. Sin embargo, una simulación no sustituye automáticamente la comprobación empírica, porque depende de supuestos, parámetros y representaciones simplificadas. Krenn et al. (2022) advierten que el valor científico de la inteligencia artificial no debe medirse solo por su rendimiento predictivo, sino por su contribución a la comprensión, explicación y formulación de conocimiento verificable.

En la fase de análisis, los modelos pueden identificar asociaciones, clasificar observaciones y generar representaciones que facilitan interpretar conjuntos de datos complejos. El riesgo aparece cuando la eficiencia técnica se confunde con solidez metodológica. Un resultado estadísticamente preciso puede carecer de validez si los datos son insuficientes, sesgados o incompatibles con la pregunta planteada. La inteligencia artificial amplía la capacidad analítica, pero no decide por sí misma qué relaciones son relevantes, qué explicaciones son aceptables ni qué grado de incertidumbre debe comunicarse según el diseño y la evidencia disponible.

La validación científica adquiere mayor importancia porque los resultados algorítmicos pueden parecer coherentes aun cuando contengan errores. Messeri y Crockett (2024) advierten que estas herramientas pueden producir ilusiones de comprensión, especialmente cuando su capacidad para generar respuestas se interpreta como conocimiento profundo del fenómeno. Por ello, validar implica contrastar resultados con evidencia independiente, revisar supuestos, documentar decisiones y examinar explicaciones alternativas. La supervisión no debe limitarse al final del proceso, sino acompañar cada fase en la que intervenga un sistema automatizado, desde la entrada de datos hasta la interpretación final.

La comunicación de resultados completa esta reconfiguración al facilitar la organización, síntesis y adaptación de contenidos científicos para diferentes públicos. Sin embargo, la responsabilidad por la exactitud, el razonamiento y la transparencia permanece en quienes firman la investigación. La inteligencia artificial redistribuye actividades, pero no reemplaza los criterios que sostienen el método científico: coherencia, contrastación, trazabilidad, revisión crítica y apertura al cuestionamiento. Su integración resulta metodológicamente valiosa cuando fortalece esos principios; se vuelve problemática cuando automatiza decisiones sin documentarlas o reduce la investigación a respuestas técnicamente convincentes.

Tabla 1

Reconfiguración de las fases del método científico mediante inteligencia artificial

Fase del método científico	Intervención de la inteligencia artificial	Aporte posible	Decisión bajo responsabilidad humana
Observación	Procesamiento de imágenes, señales, registros y datos extensos	Ampliación de la capacidad para detectar regularidades	Definir qué observar, medir y conservar
Formulación del problema	Comparación de antecedentes y detección de vacíos aparentes	Delimitación preliminar de áreas investigables	Valorar relevancia científica, social y disciplinar
Revisión de antecedentes	Organización, clasificación y síntesis documental	Identificación rápida de tendencias y debates	Comprobar calidad, pertinencia y existencia de las fuentes
Generación de hipótesis	Propuesta de relaciones y explicaciones alternativas	Ampliación del conjunto de hipótesis posibles	Evaluar plausibilidad, coherencia y comprobabilidad
Diseño metodológico	Comparación de estrategias y organización de procedimientos	Anticipación de dificultades y optimización del diseño	Seleccionar enfoque, muestra, técnicas y criterios de medición
Experimentación	Simulación, selección de alternativas y ajuste de condiciones	Reducción del espacio experimental y de los recursos requeridos	Autorizar procedimientos y validar las condiciones experimentales
Análisis	Clasificación, predicción y reconocimiento de patrones	Procesamiento de datos complejos y generación de representaciones	Determinar relevancia, validez e interpretación de los resultados
Validación	Comparación de resultados y detección de inconsistencias	Apoyo en la comprobación y revisión de supuestos	Contrastar con evidencia independiente y explicaciones alternativas
Comunicación	Organización, síntesis y adaptación del contenido	Mejora de claridad y accesibilidad de los resultados	Garantizar exactitud, transparencia y responsabilidad autoral

Nota. La tabla sintetiza cómo la inteligencia artificial puede intervenir en las fases del método científico sin sustituir las decisiones epistemológicas, metodológicas y éticas que corresponden al investigador. Elaboración propia.

La reconfiguración del método científico mediante inteligencia artificial no supone reemplazar sus fundamentos, sino ampliar las capacidades disponibles en cada fase de la investigación. Estas herramientas pueden acelerar la observación, organizar antecedentes,

generar alternativas, apoyar el análisis y mejorar la comunicación de resultados. Sin embargo, la pertinencia del problema, la elección metodológica, la interpretación de la evidencia y la validación de las conclusiones continúan bajo responsabilidad humana. Su incorporación resulta científicamente valiosa cuando fortalece la coherencia, la trazabilidad y el rigor del proceso investigativo.

1.3 Inteligencia aumentada y colaboración humano-máquina

La inteligencia aumentada describe un modelo en el que los sistemas computacionales amplían las capacidades del investigador sin sustituir su participación intelectual. A diferencia de una visión centrada en la automatización total, este enfoque considera que la inteligencia artificial resulta más valiosa cuando complementa la experiencia disciplinar, la creatividad y el juicio crítico. Goyanes y Lopezosa (2024) destacan que la integración responsable de estas tecnologías requiere colaboración humano-máquina, supervisión y pensamiento crítico. La investigación continúa siendo una actividad dirigida por personas, aunque apoyada por capacidades computacionales cada vez más amplias.

La asistencia constituye el nivel más básico de esta relación. En él, la inteligencia artificial ejecuta tareas delimitadas por el investigador, como organizar información, identificar coincidencias o generar alternativas iniciales. La persona conserva el control de los objetivos, interpreta los resultados y decide si las propuestas son útiles. Este nivel no transforma completamente la estructura de la investigación, pero reduce cargas operativas y facilita el acceso a información compleja. Su eficacia depende de que las instrucciones sean claras y de que los resultados sean revisados antes de incorporarlos al razonamiento científico.

La colaboración supone una interacción más dinámica, porque el sistema no se limita a cumplir una orden aislada, sino que participa en secuencias de exploración, comparación y ajuste. Berrios y Márquez (2025) plantean que los agentes generativos pueden apoyar procesos de aprendizaje e investigación mediante intercambios iterativos. En este modelo, el investigador formula preguntas, examina respuestas, modifica criterios y solicita nuevas aproximaciones. El conocimiento se construye mediante ciclos de interacción, pero la dirección del proceso, la selección de evidencias y la interpretación final continúan dependiendo de capacidades humanas.

La sustitución representa una situación diferente, en la que una tarea previamente realizada por personas pasa a ejecutarse de manera casi completa por un sistema. Aunque esta posibilidad puede ser adecuada en actividades repetitivas, no debe extenderse automáticamente a decisiones que exigen comprensión contextual o responsabilidad científica. Figueroa (2025) examina la inteligencia artificial como posible reemplazo, alianza o amenaza para la investigación académica. Esta distinción permite reconocer que no todas las tareas poseen el mismo nivel de complejidad ni admiten el mismo grado de delegación tecnológica.

Las fortalezas computacionales se relacionan principalmente con velocidad, escala, consistencia operativa y capacidad para procesar grandes volúmenes de datos. Los modelos pueden comparar documentos, detectar patrones y ejecutar múltiples operaciones en periodos reducidos. También pueden mantener criterios constantes cuando las instrucciones están bien definidas. Sin embargo, estas ventajas no equivalen a comprensión científica plena. Krenn et al. (2022) sostienen que el valor de la inteligencia artificial debe evaluarse por su contribución a la comprensión, no únicamente por su rendimiento predictivo o por la rapidez con la que produce resultados.

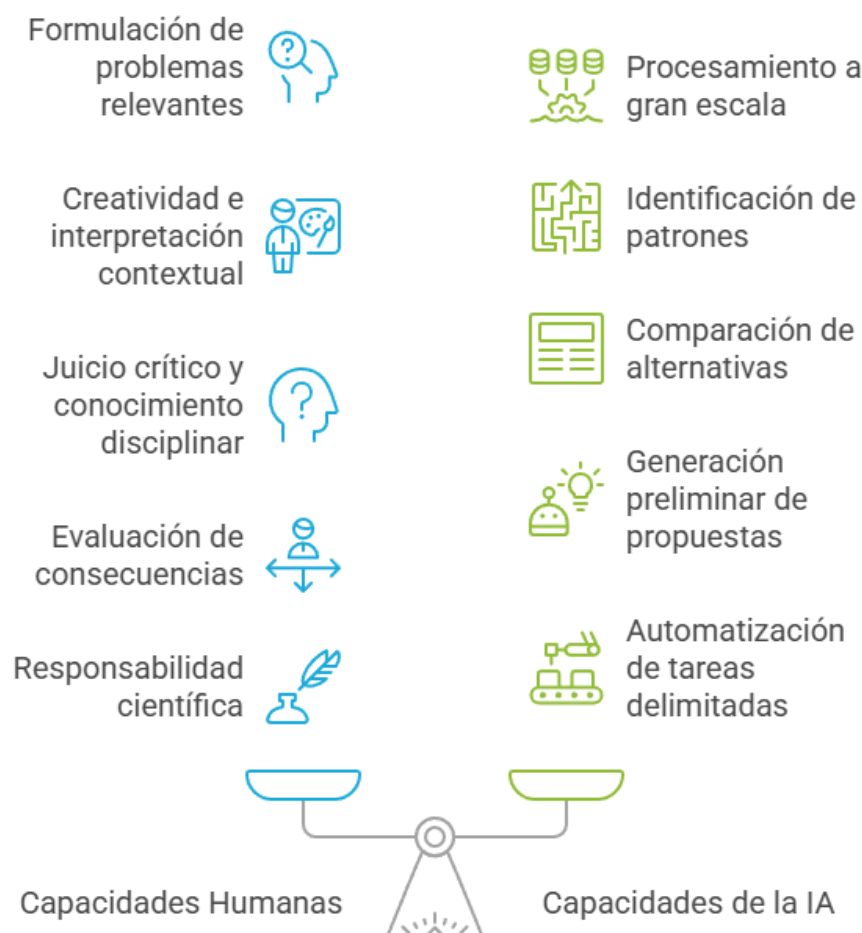
Las capacidades humanas ocupan un lugar distinto y complementario. El investigador aporta conocimiento disciplinar, sensibilidad frente al contexto, creatividad para formular problemas y capacidad para reconocer implicaciones que no aparecen directamente en los datos. También puede cuestionar supuestos, identificar consecuencias imprevistas y decidir qué resultados merecen ser estudiados. Chaaban et al. (2024) muestran que la preparación para utilizar inteligencia artificial incluye competencias técnicas, éticas y socioculturales. La colaboración efectiva requiere, por tanto, algo más que dominio instrumental: exige comprender límites, responsabilidades y condiciones de uso.

La distribución de funciones debe establecerse según la naturaleza de cada tarea. Los sistemas pueden asumir procesamiento, clasificación, generación preliminar y apoyo analítico, mientras que las personas deben conservar la definición de propósitos, la evaluación de pertinencia y la validación de conclusiones. Shao et al. (2025) proponen formas de colaboración en las que la inteligencia artificial agiliza flujos analíticos y favorece la experimentación con ideas. La utilidad de estos sistemas aumenta cuando sus funciones están delimitadas y sus resultados pueden ser examinados, corregidos o rechazados por el equipo investigador.

El control humano no consiste únicamente en aprobar una respuesta al final del proceso. Debe mantenerse desde la selección de datos y la formulación de instrucciones hasta la interpretación y comunicación de resultados. Una supervisión efectiva incluye revisar supuestos, contrastar evidencias y documentar las decisiones tomadas con apoyo tecnológico. Cuando el investigador delega sin comprender el procedimiento, puede perder capacidad para justificar sus conclusiones. La inteligencia aumentada exige una participación activa, porque la responsabilidad científica no se transfiere al sistema que produjo una recomendación, un análisis o una propuesta explicativa.

Figura 2

Convergencia de capacidades humanas y computacionales en la investigación científica



Nota. La figura representa la convergencia entre capacidades humanas y computacionales. La inteligencia artificial aporta velocidad, escala y apoyo analítico, mientras que el investigador conserva la definición de objetivos, la interpretación, la validación y la responsabilidad sobre las decisiones científicas.

La convergencia entre capacidades humanas y computacionales configura una forma de investigación en la que la inteligencia artificial amplía la velocidad, la escala y el apoyo analítico, mientras el investigador conserva la orientación científica del proceso. Esta complementariedad permite examinar más información, comparar alternativas y

automatizar tareas delimitadas sin transferir la responsabilidad sobre las decisiones. La colaboración resulta verdaderamente productiva cuando las capacidades tecnológicas se integran con juicio crítico, conocimiento disciplinar, interpretación contextual y validación humana de los resultados.

1.4 Descubrimiento científico asistido por modelos computacionales

El descubrimiento científico asistido por modelos computacionales comprende el uso de inteligencia artificial para identificar relaciones, formular expresiones matemáticas, seleccionar materiales, proponer moléculas y generar hipótesis susceptibles de comprobación. Su aporte principal no consiste en reemplazar la investigación experimental, sino en ampliar el espacio de alternativas que puede examinarse y reducirlo mediante criterios derivados de datos y conocimiento previo. Zhang et al. (2025) señalan que los modelos contemporáneos pueden intervenir desde la generación de hipótesis hasta la planificación de procedimientos orientados a contrastarlas.

En los enfoques convencionales, los investigadores seleccionan variables, proponen relaciones y ajustan modelos a partir de observaciones y fundamentos disciplinares. La inteligencia artificial introduce la posibilidad de explorar numerosas combinaciones en tiempos reducidos, descartar opciones incompatibles y localizar regularidades difíciles de reconocer manualmente. Esta capacidad resulta especialmente relevante cuando los datos presentan alta dimensionalidad o cuando existe una cantidad extensa de soluciones posibles. Sin embargo, disminuir el espacio de búsqueda no garantiza que la alternativa seleccionada sea verdadera, interpretable o científicamente relevante para el fenómeno estudiado.

AI-Descartes representa un caso de descubrimiento de ecuaciones mediante la combinación de datos experimentales y conocimiento teórico. Cornelio et al. (2023) desarrollaron un sistema orientado a encontrar fórmulas que describieran observaciones y pudieran derivarse de axiomas científicos establecidos. Esta condición diferencia su propuesta de métodos que solo buscan expresiones con buen ajuste numérico. El modelo procura que las relaciones encontradas sean compatibles con principios conocidos, lo que fortalece su interpretabilidad y permite examinar si una fórmula posee significado científico más allá de reproducir correctamente determinados datos.

No obstante, en AI-Descartes los datos y la teoría funcionan como componentes relacionados, pero mantienen cierta separación durante la búsqueda. AI-Hilbert amplía este enfoque al permitir que ambos elementos se complementen para formular relaciones científicas. Cory et al. (2024) explican que el conocimiento teórico puede compensar la escasez o el ruido de los datos, mientras que la evidencia empírica ayuda a identificar axiomas incompletos o incompatibles. Esta integración reduce el espacio de soluciones y favorece la obtención de expresiones coherentes con observaciones y fundamentos científicos.

La diferencia entre estos sistemas muestra que el descubrimiento computacional no depende exclusivamente de aumentar la cantidad de datos. La incorporación de restricciones, leyes y conocimientos previos puede orientar la búsqueda hacia resultados más significativos. AI-Hilbert fue aplicado a problemas asociados con relaciones físicas y matemáticas, demostrando que los modelos pueden utilizar información teórica para discriminar entre alternativas. Aun así, la derivación computacional de una ecuación no establece por sí misma su alcance, condiciones de aplicación o capacidad explicativa, aspectos que requieren evaluación disciplinar y contraste independiente.

En la ciencia de materiales, A-Lab evidencia cómo los modelos computacionales pueden conectarse con procesos experimentales. Szymanski et al. (2023) presentaron un laboratorio autónomo para la síntesis de materiales inorgánicos que combina cálculos, datos históricos, aprendizaje automático y aprendizaje activo. El sistema selecciona objetivos y procedimientos experimentales, mientras una infraestructura automatizada ejecuta y analiza las síntesis. Su importancia radica en enlazar predicciones computacionales con resultados físicos, reduciendo la distancia entre la identificación teórica de materiales prometedores y su obtención efectiva en el laboratorio.

El aprendizaje activo empleado en estos entornos permite que los resultados de cada experimento orienten la selección del siguiente. De este modo, el sistema no recorre indiscriminadamente todas las combinaciones posibles, sino que actualiza sus decisiones según la evidencia obtenida. Esta estrategia puede reducir materiales, tiempo y operaciones fallidas, aunque depende de la calidad de las mediciones y de los criterios utilizados para definir el objetivo experimental. Los resultados negativos también adquieren valor, porque permiten modificar las estimaciones y descartar regiones poco prometedoras del espacio de búsqueda.

En química, Boiko et al. (2023) desarrollaron Coscientist, un sistema basado en un gran modelo de lenguaje conectado con búsqueda documental, ejecución de código y automatización experimental. El sistema fue diseñado para planificar y realizar procedimientos químicos complejos mediante la coordinación de diferentes herramientas. Este caso demuestra que un modelo generativo puede desempeñar funciones que exceden la producción textual cuando dispone de acceso controlado a fuentes, programas e instrumentos. Sin embargo, su capacidad depende de la precisión de las instrucciones, la integración técnica y los mecanismos de seguridad establecidos.

Los modelos también pueden contribuir a formular y priorizar hipótesis en campos donde el conocimiento se encuentra distribuido entre numerosas publicaciones. Gottweis et al. (2026) presentaron Co-Scientist, una arquitectura multiagente que genera, debate, clasifica y modifica propuestas científicas mediante procesos iterativos. El sistema fue aplicado a problemas biomédicos relacionados con reposicionamiento de fármacos, fibrosis hepática y transferencia genética bacteriana. Las hipótesis seleccionadas fueron sometidas posteriormente a experimentos independientes, lo que permitió distinguir entre propuestas computacionalmente plausibles y resultados respaldados por evidencia de laboratorio.

Estos casos muestran distintos niveles de participación de la inteligencia artificial en el descubrimiento. AI-Descartes y AI-Hilbert se concentran en relaciones matemáticas sustentadas por datos y teoría; A-Lab integra selección computacional y síntesis de materiales; Coscientist coordina información, código y experimentación química; y Co-Scientist organiza agentes para producir hipótesis biomédicas. En todos ellos, el modelo funciona como parte de un sistema más amplio de investigación. Su rendimiento no puede evaluarse únicamente por la novedad de sus propuestas, sino por su coherencia, trazabilidad y validación independiente.

Por tanto, debe diferenciarse cuidadosamente la sugerencia computacional de la comprobación científica. Un modelo puede identificar una correlación, producir una ecuación o proponer un material, pero el hallazgo adquiere validez cuando resiste pruebas empíricas, comparación con explicaciones alternativas y revisión especializada. Los sistemas computacionales aceleran la exploración y permiten administrar espacios de búsqueda difíciles de abordar manualmente; no eliminan la necesidad de justificar procedimientos y reconocer incertidumbres. El descubrimiento asistido conserva una

estructura colaborativa en la que cálculo, teoría, experimentación e interpretación cumplen funciones complementarias.

Tabla 2

Casos de descubrimiento científico asistido por modelos computacionales

Sistema o estudio	Área científica	Información integrada	Aporte principal	Forma de comprobación
AI-Descartes	Física y modelización científica	Datos experimentales y axiomas teóricos	Descubrimiento de ecuaciones ajustadas a datos y derivables desde conocimiento previo	Análisis de ajuste, derivación teórica y comparación con relaciones conocidas
AI-Hilbert	Física y matemáticas	Datos, axiomas y optimización algebraica	Formulación de relaciones mediante interacción entre evidencia y teoría	Consistencia con datos, derivabilidad y evaluación de axiomas
A-Lab	Ciencia de materiales	Cálculos, literatura, datos históricos y resultados experimentales	Selección y síntesis automatizada de materiales inorgánicos	Caracterización física de los productos obtenidos
Coscientist	Química	Literatura, documentación técnica, código e instrumentación	Planificación y ejecución coordinada de experimentos químicos	Ejecución de procedimientos y comprobación de productos experimentales
Co-Scientist	Biomedicina	Literatura, agentes especializados, criterios científicos y retroalimentación humana	Generación, debate, clasificación y refinamiento de hipótesis	Validación independiente mediante experimentos de laboratorio

Nota. La tabla compara sistemas que intervienen en diferentes componentes del descubrimiento científico. La generación computacional de relaciones, materiales o hipótesis constituye una etapa de exploración que requiere comprobación teórica, empírica o experimental.

Los casos presentados muestran que el descubrimiento asistido por modelos computacionales puede adoptar formas distintas según el campo científico y el tipo de evidencia disponible. Algunos sistemas se orientan a formular ecuaciones, otros a seleccionar materiales, coordinar experimentos o generar hipótesis biomédicas. Aunque estas herramientas amplían la capacidad de exploración y reducen el espacio de búsqueda, sus propuestas solo adquieren valor científico cuando son contrastadas mediante derivación teórica, análisis de consistencia, caracterización experimental o validación independiente.

Figura 3

Proceso general del descubrimiento científico asistido por modelos computacionales



Nota. La figura representa la secuencia mediante la cual una propuesta generada por inteligencia artificial pasa de la exploración computacional a la comprobación científica.

El descubrimiento científico asistido por modelos computacionales constituye un proceso en el que la inteligencia artificial contribuye a explorar relaciones, formular propuestas y orientar nuevas vías de análisis a partir de datos y conocimiento previo. Sin embargo, la utilidad de estas herramientas no reside únicamente en generar resultados plausibles, sino en su capacidad para integrarse con procedimientos de comprobación teórica, empírica o experimental.

1.5 Transformaciones disciplinares y convergencia científica

La inteligencia artificial no transforma todas las disciplinas de la misma manera, porque cada campo posee objetos de estudio, datos, criterios de validación y tradiciones metodológicas particulares. En algunas áreas, su incorporación se concentra en analizar grandes volúmenes de información; en otras, permite automatizar experimentos, integrar

modalidades o explorar soluciones difíciles de abordar manualmente. Rodríguez et al. (2025) muestran que las técnicas de inteligencia artificial se extienden hacia diversos ámbitos científicos, aunque su adopción depende de la disponibilidad de datos, infraestructura y competencias especializadas.

En las ciencias de la salud, la transformación se relaciona con la integración de información clínica, imágenes, datos moleculares y antecedentes de pacientes. Los modelos pueden apoyar la clasificación de riesgos, la identificación de patrones diagnósticos y la selección de tratamientos adaptados a características individuales. Truhn et al. (2024) señalan que los modelos fundacionales multimodales permiten combinar fuentes heterogéneas en oncología de precisión. Sin embargo, su validación exige comparar resultados con evidencia clínica, evaluar poblaciones diversas y considerar consecuencias directas sobre decisiones que afectan la salud y seguridad de las personas.

En química, la inteligencia artificial interviene principalmente en la exploración de moléculas, la predicción de propiedades y la planificación de procedimientos experimentales. Los modelos reducen el número de combinaciones que deben evaluarse físicamente y ayudan a seleccionar candidatos con mayor probabilidad de cumplir determinados objetivos. Esta transformación modifica la relación entre diseño y experimentación, porque parte del espacio químico puede examinarse computacionalmente antes de ejecutar pruebas. No obstante, una molécula prometedora según un modelo necesita síntesis, caracterización y comprobación experimental para determinar su comportamiento en condiciones reales.

La ciencia de materiales comparte algunas características con la química, pero su transformación se orienta hacia la búsqueda de composiciones, estructuras y procesos capaces de producir propiedades específicas. Tom et al. (2024) describen cómo los laboratorios autónomos combinan aprendizaje automático, instrumentación y automatización para acelerar la experimentación en química y materiales. En estos campos, la validación depende de resultados físicos reproducibles, como estabilidad, resistencia o conductividad. La IA no sustituye las mediciones, sino que reorganiza la selección de candidatos y permite aprender de experimentos anteriores para decidir cuáles conviene ejecutar después.

En ingeniería, los modelos computacionales contribuyen a optimizar diseños, predecir fallas, simular condiciones y ajustar procesos productivos. La transformación

metodológica se observa en la combinación de datos provenientes de sensores, modelos físicos y registros operativos. Esta integración facilita comparar alternativas sin construir inmediatamente todos los prototipos posibles. Sin embargo, las soluciones deben cumplir restricciones técnicas, económicas, ambientales y de seguridad. Un diseño optimizado matemáticamente puede resultar inviable si no considera materiales disponibles, mantenimiento, normas de funcionamiento o condiciones específicas del entorno donde será implementado.

En educación, la inteligencia artificial se utiliza para analizar trayectorias de aprendizaje, identificar necesidades, organizar información académica y apoyar procesos de investigación pedagógica. Casanova y Martínez (2024) evidencian el crecimiento de la producción científica sobre inteligencia artificial y educación, lo que refleja una rápida consolidación temática. Molina (2024) destaca factores que condicionan su uso en investigación educativa. A diferencia de las ciencias experimentales, la validación no puede limitarse a precisión técnica, porque también debe considerar contexto escolar, diversidad estudiantil, prácticas docentes y significado pedagógico de los resultados.

Las ciencias sociales incorporan inteligencia artificial para examinar discursos, documentos, redes, comportamientos y grandes conjuntos de información producida en entornos digitales. Estas capacidades amplían la escala del análisis, pero introducen tensiones relacionadas con la interpretación contextual. Una frecuencia elevada, una asociación estadística o una categoría generada automáticamente no explican por sí mismas las prácticas y experiencias humanas. Por ello, la transformación disciplinar depende de combinar procedimientos computacionales con marcos teóricos, lectura crítica y conocimiento histórico. La eficiencia analítica debe mantenerse subordinada a preguntas socialmente relevantes y metodológicamente justificadas.

Las diferencias disciplinares también se manifiestan en los criterios de validación. En salud se requiere evidencia clínica y evaluación de seguridad; en química y materiales, comprobación experimental; en ingeniería, desempeño bajo condiciones operativas; en educación, pertinencia pedagógica; y en ciencias sociales, coherencia interpretativa y contextual. Zhang et al. (2025) sostienen que los modelos de lenguaje pueden intervenir en distintas etapas de la investigación, pero su utilidad depende del campo y del problema. En consecuencia, no existe un único procedimiento para validar resultados asistidos por inteligencia artificial en todas las ciencias.

La convergencia científica surge cuando problemas complejos requieren integrar conocimientos, datos y métodos procedentes de varias disciplinas. Ocampo et al. (2024) vinculan tecnologías convergentes, inteligencia artificial y neurociencias con la formación de investigadores capaces de trabajar entre campos. Esta convergencia puede reunir, por ejemplo, medicina, biología, informática y estadística, o combinar ingeniería, ciencias ambientales y análisis de datos. La colaboración interdisciplinaria permite formular preguntas más amplias, aunque también exige acordar vocabularios, estándares, criterios de calidad y responsabilidades entre equipos con tradiciones metodológicas diferentes.

Por tanto, la principal transformación no consiste únicamente en introducir algoritmos dentro de cada disciplina, sino en reorganizar la producción del conocimiento y las relaciones entre campos científicos. La inteligencia artificial facilita compartir datos, conectar modelos y trasladar técnicas entre áreas, pero sus resultados deben interpretarse según las exigencias de cada dominio. La convergencia es productiva cuando respeta la especificidad disciplinar y establece mecanismos comunes de validación. Sin esta articulación, existe el riesgo de aplicar soluciones técnicamente sofisticadas que no respondan adecuadamente a los problemas científicos estudiados.

Tabla 3
Transformaciones disciplinares asociadas con el uso de inteligencia artificial

Disciplina	Datos utilizados	Aplicación de la inteligencia artificial	Transformación metodológica	Limitación principal
Salud	Historias clínicas, imágenes, datos moleculares y señales biológicas	Clasificación de riesgos, apoyo diagnóstico e integración multimodal	Combinación de evidencias clínicas y computacionales para decisiones personalizadas	Sesgos poblacionales, privacidad y necesidad de validación clínica
Química	Estructuras moleculares, propiedades, literatura y resultados experimentales	Diseño de moléculas, predicción de propiedades y planificación experimental	Reducción del espacio químico antes de la síntesis	Diferencia entre predicción y comportamiento experimental real
Ciencia de materiales	Composiciones, estructuras, propiedades y	Selección de materiales y optimización de experimentos	Integración de modelos, aprendizaje activo	Infraestructura costosa y dificultad para

Disciplina	Datos utilizados	Aplicación de la inteligencia artificial	Transformación metodológica	Limitación principal
Ingeniería	mediciones físicas	Optimización, predicción de fallas y evaluación de diseños	y experimentación automatizada	generalizar resultados
	Sensores, simulaciones, registros operativos y modelos físicos		Comparación computacional de alternativas antes de construir prototipos	Restricciones técnicas, económicas, ambientales y de seguridad
Educación	Registros académicos, interacciones, evaluaciones y documentos	Análisis de aprendizaje y apoyo a la investigación pedagógica	Ampliación del análisis de trayectorias y contextos educativos	Riesgo de reducir procesos formativos a indicadores cuantificables
Ciencias sociales	Textos, redes, encuestas, archivos y datos digitales	Análisis de discursos, comportamientos y relaciones sociales	Procesamiento de información a gran escala combinado con interpretación teórica	Pérdida de contexto, simplificación y sesgos culturales

Nota. La tabla compara la transformación metodológica producida por la inteligencia artificial en distintas disciplinas. Las aplicaciones y criterios de validación varían según la naturaleza de los datos, los objetivos científicos y las consecuencias de los resultados.

La inteligencia artificial transforma las disciplinas científicas de acuerdo con la naturaleza de sus datos, métodos y criterios de validación. En salud, química, materiales e ingeniería fortalece la predicción, el diseño y la experimentación, mientras que en educación y ciencias sociales amplía el análisis de trayectorias, discursos y contextos. Estas posibilidades no son equivalentes entre campos ni están exentas de limitaciones. Su incorporación resulta pertinente cuando las herramientas se adaptan a las exigencias disciplinares y sus resultados se validan mediante procedimientos científicos específicos.

1.6 Infraestructuras digitales para la investigación inteligente

La investigación inteligente no depende únicamente de modelos avanzados, sino de un conjunto de infraestructuras capaces de almacenar, conectar, procesar y documentar información científica. Estas estructuras incluyen repositorios, bases de datos, capacidad de cómputo, plataformas colaborativas, interfaces de programación, sistemas de control de versiones e instrumentación automatizada. Su integración permite que los datos circulen entre distintas etapas del trabajo investigador. Sin una infraestructura organizada, la inteligencia artificial opera sobre recursos fragmentados, difíciles de verificar y con escasas posibilidades de reutilización en nuevos proyectos.

Los repositorios constituyen una capa fundamental porque conservan publicaciones, conjuntos de datos, protocolos, códigos y resultados de investigación. Su utilidad no se limita al almacenamiento: también facilitan la identificación, recuperación, citación y reutilización de los recursos depositados. Choque et al. (2025) relacionan la ciencia abierta con la transparencia, el acceso a datos y la colaboración interdisciplinaria. Para que estos espacios apoyen sistemas inteligentes, deben incorporar metadatos comprensibles, formatos consistentes y criterios que permitan reconocer procedencia, condiciones de uso y modificaciones realizadas.

Los sistemas de datos conectan los repositorios con procesos de preparación, análisis y aprendizaje automático. Esta infraestructura comprende mecanismos de captura, limpieza, normalización, integración y actualización de información procedente de diversas fuentes. Su calidad depende de que cada transformación pueda rastrearse y de que las variables conserven un significado estable. Cuando los datos carecen de documentación, los modelos pueden procesarlos técnicamente, pero resulta difícil interpretar sus resultados. Por ello, la organización de datos debe atender tanto a su estructura computacional como a su sentido científico y disciplinar.

La capacidad de cómputo constituye otra condición decisiva, especialmente cuando se procesan imágenes, secuencias, simulaciones o grandes colecciones documentales. Los recursos pueden distribuirse entre servidores institucionales, infraestructuras de alto rendimiento y entornos de computación en la nube. Esta diversidad permite ajustar la capacidad tecnológica a las características del proyecto. Sin embargo, también genera desigualdades entre instituciones y regiones. El acceso limitado a procesamiento, almacenamiento y mantenimiento puede restringir la participación de determinados equipos, incluso cuando poseen problemas relevantes, datos valiosos y conocimientos disciplinares suficientes.

Las interfaces de programación de aplicaciones permiten que plataformas, modelos y bases de datos intercambien información mediante procedimientos definidos. A través de estas conexiones, un sistema puede consultar repositorios, ejecutar análisis, recuperar documentos o enviar instrucciones a otros componentes. Las interfaces favorecen la modularidad porque evitan construir todo el entorno desde cero. No obstante, su uso exige controlar permisos, formatos, límites de acceso y posibles cambios en los servicios conectados. Una investigación dependiente de interfaces externas debe documentar cómo fueron utilizadas para conservar la trazabilidad del procedimiento.

La gestión de versiones permite registrar modificaciones en datos, código, instrucciones y documentos elaborados durante el proyecto. Van et al. (2024) proponen mediante ENCORE una estructura estandarizada para organizar y documentar los componentes de una investigación computacional. Este enfoque integra archivos, análisis y resultados dentro de una organización que facilita comprender cómo se produjo cada salida. La gestión de versiones no debe considerarse una práctica exclusiva de programadores, porque también permite reconstruir decisiones, comparar alternativas y evitar que los cambios realizados durante el estudio queden ocultos.

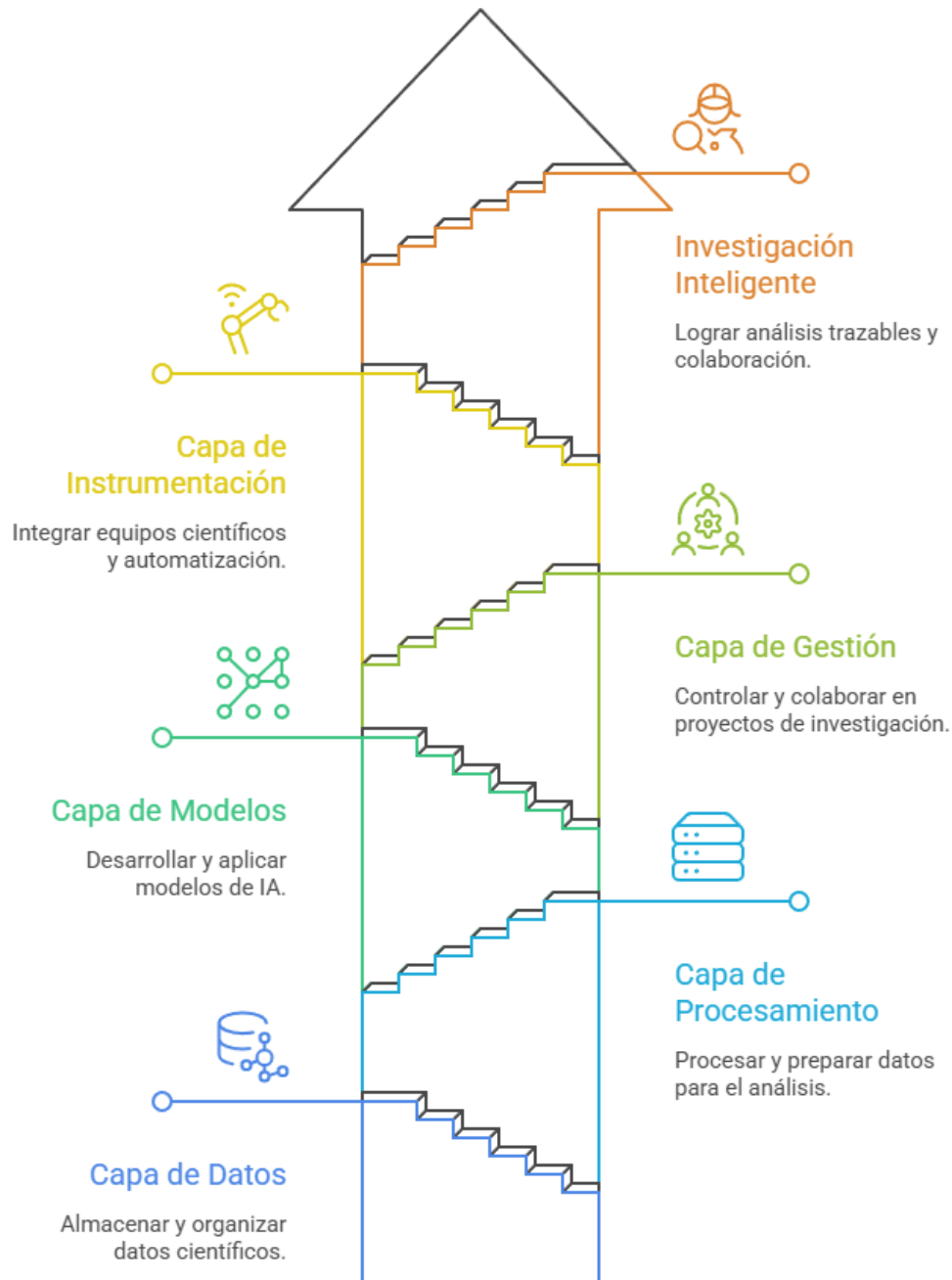
Las plataformas colaborativas complementan esta organización al permitir que equipos distribuidos compartan documentos, datos, procedimientos y observaciones. La cultura digital institucional influye en la forma en que estos recursos son adoptados y utilizados. Henríquez et al. (2025) muestran que el empleo académico de tecnologías se relaciona con el manejo de información, la comunicación y las políticas institucionales que favorecen su integración. Una plataforma resulta útil cuando establece responsabilidades, permisos y prácticas comunes; por sí sola no garantiza cooperación efectiva, calidad documental ni continuidad del conocimiento producido.

En los campos experimentales, la infraestructura digital se extiende hacia sensores, instrumentos conectados y sistemas robóticos. Tom et al. (2024) explican que los laboratorios autónomos articulan hardware, software, automatización de procedimientos y planificación computacional. La instrumentación produce datos que pueden alimentar modelos y orientar nuevas operaciones, estableciendo una relación continua entre espacio físico y entorno digital. En este apartado, lo relevante no es el nivel de autonomía alcanzado, sino reconocer que la investigación inteligente requiere compatibilidad entre equipos, protocolos, datos y sistemas de control.

La integración de estas capas permite construir entornos en los que los investigadores consultan datos, ejecutan procedimientos y documentan resultados sin perder la conexión entre sus componentes. Shao et al. (2025) muestran que un colaborador basado en inteligencia artificial puede automatizar flujos analíticos, favorecer la iteración y facilitar la reproducibilidad cuando opera sobre recursos estructurados. La infraestructura, por tanto, no es un soporte secundario, sino la base que determina qué puede automatizarse, verificarse y compartirse. Su diseño debe equilibrar interoperabilidad, seguridad, acceso, sostenibilidad y control científico.

Figura 4

Arquitectura digital para la investigación inteligente



Nota. La figura representa las principales capas que sostienen la investigación inteligente. Su funcionamiento depende de la interoperabilidad entre datos, capacidad computacional, modelos, sistemas de gestión e instrumentación científica.

La investigación inteligente depende de una arquitectura digital articulada en la que datos, capacidad de cómputo, modelos, sistemas de gestión e instrumentación científica funcionan de manera integrada. Cada capa cumple una función específica, pero su valor surge de la interoperabilidad y la trazabilidad entre procesos. Esta organización permite conectar actividades digitales y experimentales, favorecer la colaboración y reutilizar recursos científicos. Sin embargo, su eficacia requiere documentación, control de

versiones, seguridad y condiciones técnicas que garanticen la continuidad y confiabilidad del entorno investigativo.

1.7 Competencias del investigador en entornos asistidos por inteligencia artificial

La incorporación de inteligencia artificial en la investigación exige competencias que superan el manejo operativo de una plataforma. Un investigador competente debe comprender qué puede hacer el sistema, bajo qué condiciones produce resultados y cuáles decisiones no pueden delegarse. Chaaban et al. (2024) plantean que la preparación para trabajar con IA integra capacidades técnicas, evaluación crítica y criterios éticos dentro de un contexto sociocultural. Por ello, la formación debe orientarse a supervisar procesos, justificar elecciones y reconocer cuándo una respuesta computacional carece de sustento suficiente.

La alfabetización en inteligencia artificial constituye la competencia inicial. Incluye comprender nociones básicas sobre aprendizaje automático, modelos generativos, datos de entrenamiento, probabilidades y limitaciones de las salidas. No se requiere que todos los investigadores sean especialistas informáticos, pero sí que puedan distinguir entre una predicción, una clasificación, una generación textual y una explicación científica. Krenn et al. (2022) sostienen que el valor de la IA para la ciencia no debe evaluarse únicamente mediante rendimiento, sino también por su aporte a la comprensión del fenómeno investigado.

El manejo de datos representa una segunda competencia esencial, porque la calidad de los resultados depende de la procedencia, estructura y representatividad de la información utilizada. El investigador debe reconocer datos incompletos, variables mal definidas, registros duplicados y posibles sesgos de selección. También necesita comprender procesos de limpieza, normalización y documentación. Vimos et al. (2024) muestran que el uso académico de inteligencia artificial exige capacidades para integrar recursos digitales en las distintas fases investigativas, sin separar el desempeño tecnológico de la calidad metodológica del estudio.

La comprensión funcional de los modelos permite interpretar sus resultados sin atribuirles capacidades que no poseen. Esto supone conocer qué entradas reciben, qué tipo de salida generan y cómo pueden variar sus respuestas según instrucciones, versiones o configuraciones. Messeri y Crockett (2024) advierten que los sistemas de IA pueden

producir una ilusión de comprensión cuando sus respuestas convincentes se confunden con explicaciones científicas verificadas. La competencia investigativa requiere examinar supuestos, incertidumbre y límites de generalización antes de convertir una salida algorítmica en evidencia o argumento.

La formulación de instrucciones constituye otra habilidad relevante, especialmente en modelos generativos. Diseñar una indicación adecuada implica establecer propósito, contexto, restricciones, formato de respuesta y criterios de calidad. Sin embargo, escribir instrucciones detalladas no garantiza resultados correctos. Lalaleo et al. (2024) identifican la necesidad de fortalecer capacidades para utilizar la IA en búsqueda, organización y análisis de información científica. La instrucción debe entenderse como parte de un proceso iterativo que incluye revisar, contrastar, reformular y documentar las respuestas obtenidas.

La programación básica amplía la autonomía del investigador al permitirle comprender y adaptar procedimientos de análisis. Esta competencia puede incluir lectura de código, identificación de errores, ejecución de scripts y revisión de resultados producidos automáticamente. No todas las disciplinas requieren el mismo nivel técnico, pero depender completamente de instrucciones generadas por terceros limita la capacidad de auditoría. Shao et al. (2025) muestran que los sistemas colaborativos pueden automatizar flujos analíticos; aun así, el investigador necesita entender las operaciones realizadas para evaluar su pertinencia y reproducibilidad.

La validación constituye el núcleo que conecta las competencias técnicas con el rigor científico. El investigador debe comparar resultados con fuentes primarias, métodos alternativos, datos independientes o criterios expertos. Una salida no se valida porque esté bien redactada ni porque coincida con una expectativa inicial. Cedeño et al. (2024) destacan que la investigación universitaria con inteligencia artificial demanda formación para utilizar estas tecnologías de manera crítica y responsable. Validar también implica registrar qué resultados fueron rechazados, corregidos o modificados durante el proceso.

El pensamiento crítico permite decidir cuándo la inteligencia artificial aporta valor y cuándo introduce complejidad innecesaria. Esta competencia comprende cuestionar supuestos, identificar contradicciones, reconocer asociaciones espurias y evaluar si una herramienta responde realmente a la pregunta planteada. Chaaban et al. (2024) encontraron perspectivas diversas sobre la preparación de los investigadores, algunas

centradas en habilidades técnicas y otras en consideraciones éticas o colaborativas. Esta diversidad indica que la competencia no es única ni lineal, sino una combinación de conocimientos, actitudes y prácticas ajustadas al contexto disciplinar.

La dimensión ética incluye privacidad, consentimiento, seguridad de los datos y evaluación de consecuencias. El investigador debe identificar cuándo la información no puede incorporarse a servicios externos, qué restricciones acompañan su reutilización y cómo evitar exposiciones innecesarias. También necesita reconocer sesgos que puedan afectar a poblaciones o contextos específicos. Las tecnologías convergentes demandan habilidades digitales y socioemocionales, además de un uso responsable de la inteligencia artificial durante la formación científica (Ocampo et al., 2024).

Tabla 4
Competencias del investigador en entornos asistidos por inteligencia artificial

Competencia	Descripción	Aplicación investigativa	Riesgo de su ausencia	Estrategia formativa
Alfabetización en IA	Comprensión básica de modelos, datos, probabilidades y límites	Seleccionar sistemas adecuados y reconocer el alcance de sus resultados	Sobreestimación de capacidades y uso inadecuado	Cursos introductorios con análisis de casos científicos
Manejo de datos	Evaluación de procedencia, calidad, estructura y representatividad	Preparar insumos pertinentes para análisis y modelado	Resultados sesgados, incompletos o difíciles de interpretar	Talleres prácticos de limpieza, documentación y control de calidad
Comprensión de modelos	Conocimiento funcional de entradas, procesos generales y salidas	Interpretar resultados según las condiciones del modelo	Confusión entre predicción, generación y explicación	Comparación guiada entre modelos y análisis de limitaciones
Formulación de instrucciones	Definición de propósito, contexto, restricciones y formato	Obtener respuestas orientadas a una tarea científica delimitada	Salidas ambiguas, irrelevantes o difíciles de comprobar	Ejercicios iterativos de diseño, revisión y documentación
Programación básica	Lectura, ejecución y revisión	Automatizar análisis, adaptar procedimientos	Dependencia técnica y escasa	Proyectos aplicados con lenguajes y

Competencia	Descripción	Aplicación investigativa	Riesgo de su ausencia	Estrategia formativa
	elemental de código	y detectar errores	capacidad de auditoría	entornos reproducibles
Validación	Contraste de salidas con evidencia, métodos y criterios independientes	Determinar si un resultado puede incorporarse al estudio	Aceptación de errores plausibles como evidencia válida	Protocolos de triangulación y ejercicios de verificación
Pensamiento crítico	Evaluación de supuestos, coherencia, pertinencia y alternativas	Decidir cuándo usar, corregir o descartar una salida	Automatización acrítica y pérdida de control metodológico	Discusión de casos, errores y decisiones controvertidas
Ética y privacidad	Protección de datos, consentimiento, seguridad y evaluación de consecuencias	Usar información y sistemas de manera responsable	Exposición de datos, afectación de participantes y daños previsibles	Formación mediante escenarios de riesgo y dilemas aplicados
Comunicación científica	Explicación clara de procedimientos, resultados y limitaciones	Presentar investigaciones comprensibles, auditables y prudentes	Opacidad metodológica y exageración del aporte tecnológico	Escritura científica, revisión entre pares y divulgación responsable

Nota. La tabla organiza las competencias necesarias para integrar inteligencia artificial en la investigación con rigor. El dominio instrumental constituye solo una parte de la preparación; su uso científico requiere validación, comprensión metodológica, responsabilidad ética y comunicación transparente.

La incorporación rigurosa de inteligencia artificial en la investigación depende de un conjunto articulado de competencias técnicas, metodológicas y éticas. La alfabetización en IA, el manejo de datos y la comprensión de modelos permiten utilizar las herramientas con mayor criterio, mientras la validación y el pensamiento crítico evitan aceptar resultados plausibles sin evidencia suficiente. A ello se suman la programación básica, la protección de la privacidad y la comunicación transparente. La formación del investigador debe orientarse, por tanto, a conservar autonomía intelectual, capacidad de auditoría y responsabilidad científica.

Capítulo

2

Generación del conocimiento con inteligencia artificial

Ariel Oswaldo Fernández Loor



2.1 Exploración automatizada de literatura científica

La exploración de literatura científica constituye una fase orientada a comprender cómo se organiza un campo, cuáles debates predominan y qué áreas permanecen insuficientemente estudiadas. La inteligencia artificial amplía este proceso al permitir consultas sobre colecciones extensas, clasificación de documentos y generación de panoramas preliminares. Goyanes y Lopezosa (2024) utilizaron una revisión exploratoria para identificar tendencias temáticas relacionadas con ChatGPT y los métodos de investigación. Su experiencia muestra que la automatización puede apoyar el reconocimiento inicial del terreno científico, siempre que los resultados sean depurados y examinados críticamente.

La búsqueda automatizada comienza con la formulación de una consulta suficientemente delimitada para recuperar publicaciones pertinentes. Los modelos pueden ayudar a proponer términos relacionados, equivalencias conceptuales y combinaciones entre palabras clave, especialmente cuando el campo utiliza vocabularios diversos. Este apoyo permite ampliar la recuperación y descubrir expresiones que el investigador podría no haber considerado inicialmente. Sin embargo, la definición del tema, el periodo, las disciplinas y los tipos documentales continúa bajo control humano, porque una estrategia excesivamente amplia produce ruido y una demasiado restrictiva puede excluir contribuciones relevantes.

Después de la consulta, los sistemas pueden recuperar documentos desde bases bibliográficas, repositorios y colecciones institucionales. La automatización facilita reunir títulos, resúmenes, palabras clave, autores y metadatos en conjuntos analizables. Esta etapa exige distinguir entre acceso al registro bibliográfico y acceso al texto completo, porque una referencia recuperada no garantiza que su contenido pueda evaluarse. También es necesario revisar duplicados, versiones preliminares y documentos sin información suficiente. La calidad de la exploración depende tanto de la amplitud de las fuentes consultadas como de la consistencia de sus registros.

La clasificación automatizada permite organizar publicaciones según temas, métodos, disciplinas o problemas científicos. Casanova y Martínez (2024) emplearon análisis cuantitativo para examinar la producción sobre inteligencia artificial y educación, mostrando cómo los datos bibliográficos permiten reconocer evolución temática, concentración de publicaciones y relaciones dentro de un campo. Estas técnicas son útiles

para construir una visión estructurada, aunque las categorías obtenidas deben interpretarse con cautela. Una agrupación estadística refleja semejanzas documentales, pero no sustituye la lectura conceptual de los estudios incluidos.

El agrupamiento temático puede revelar conjuntos de investigación, conceptos emergentes y relaciones entre áreas antes consideradas separadas. También permite elaborar mapas de coocurrencia, redes de autores y conexiones entre referencias citadas. Estas representaciones ayudan a identificar núcleos consolidados, líneas periféricas y posibles convergencias interdisciplinarias. Shao et al. (2025) muestran que los sistemas colaborativos de inteligencia artificial pueden apoyar flujos analíticos e iterativos en la ciencia de la ciencia. No obstante, las redes bibliométricas deben entenderse como representaciones parciales condicionadas por la cobertura de las bases consultadas.

La identificación de tendencias se apoya en la frecuencia, evolución temporal y conexión entre conceptos. A partir de estos indicadores, la inteligencia artificial puede señalar temas en crecimiento, enfoques dominantes y áreas que pierden presencia. Sin embargo, una tendencia cuantitativa no equivale necesariamente a una prioridad científica. El número de publicaciones puede estar influido por financiamiento, disponibilidad de datos, interés editorial o concentración geográfica. Por ello, los resultados deben contrastarse con la relevancia del problema, la calidad de los estudios y las necesidades reales de conocimiento dentro del campo analizado.

La detección de vacíos requiere todavía mayor cautela. La ausencia aparente de publicaciones puede indicar un tema poco estudiado, pero también una estrategia de búsqueda incompleta, diferencias terminológicas o exclusión de literatura regional. Ghareeb et al. (2026) muestran que los sistemas multiagente pueden integrar exploración documental, generación de propuestas y revisión iterativa. Aplicada a la literatura, esta lógica permite contrastar consultas, comparar resultados y reformular búsquedas. El vacío adquiere valor científico cuando ha sido comprobado mediante fuentes diversas y no únicamente inferido por un modelo.

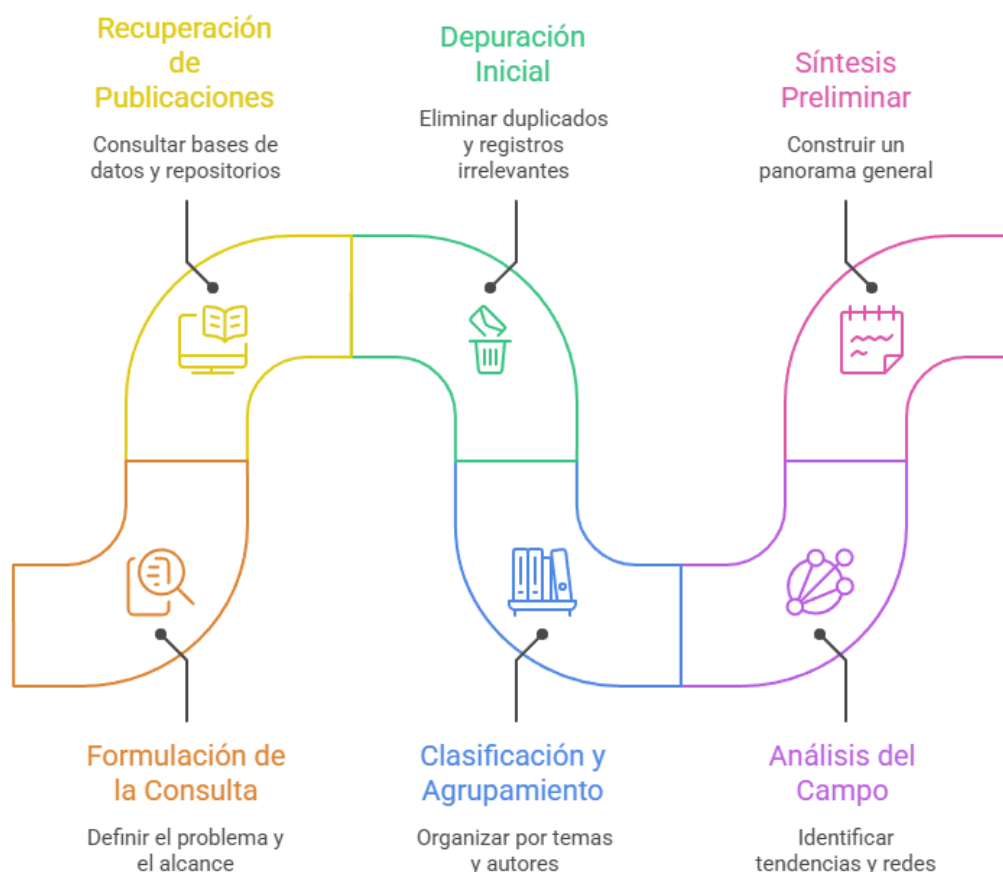
El cribado automatizado ayuda a revisar títulos y resúmenes para separar documentos potencialmente pertinentes de aquellos alejados del objetivo. Li et al. (2024) compararon grandes modelos de lenguaje en tareas de evaluación de resúmenes y observaron que pueden aplicar criterios definidos por los investigadores, aunque su rendimiento varía según el modelo y la formulación utilizada. Syriani et al. (2024) también evaluaron el

empleo de ChatGPT en el cribado documental. Estos resultados respaldan un uso asistido, pero no justifican aceptar decisiones automáticas sin comprobación humana.

La síntesis inicial integra temas, tendencias, autores, métodos y vacíos en una representación comprensible del campo. Esta exploración se diferencia de una revisión sistemática porque busca orientar, cartografiar y generar preguntas, sin desarrollar necesariamente un protocolo exhaustivo de selección y extracción. Antes de utilizar sus resultados, el investigador debe verificar títulos, autores, años, DOI y contenido de las fuentes originales. La exploración automatizada es valiosa cuando reduce desorden y amplía perspectivas; pierde rigor cuando sustituye la lectura, presenta referencias inexistentes o convierte clasificaciones preliminares en conclusiones definitivas.

Figura 5

Flujo de exploración automatizada de literatura científica



Nota. La figura representa un proceso de exploración bibliográfica asistido por inteligencia artificial. Las etapas automatizadas permiten ampliar y organizar la literatura, mientras que la revisión humana comprueba la existencia, pertinencia, calidad y significado de las publicaciones recuperadas.

La exploración automatizada de literatura científica permite ampliar y organizar la búsqueda documental mediante procesos de recuperación, depuración, clasificación y análisis temático. Estas funciones facilitan la identificación de tendencias, relaciones y

vacíos aparentes, pero no sustituyen la evaluación crítica de las fuentes. La revisión humana continúa siendo indispensable para comprobar la existencia de las publicaciones, valorar su pertinencia y evitar interpretaciones erróneas. La automatización resulta útil cuando orienta el conocimiento del campo sin reemplazar la lectura, la verificación y el juicio científico.

2.2 Generación y priorización de hipótesis

La formulación de hipótesis constituye un punto de enlace entre el conocimiento disponible y la producción de explicaciones susceptibles de comprobación. La inteligencia artificial amplía esta fase al relacionar literatura, datos, modelos teóricos y resultados previos para proponer alternativas que podrían no resultar evidentes mediante una revisión manual. Suazo (2023) reconoce la generación y validación de hipótesis como una de las aplicaciones relevantes de la IA en investigación. Sin embargo, una propuesta generada computacionalmente solo representa una posibilidad inicial y no una afirmación científica confirmada.

Las hipótesis pueden originarse en observaciones inesperadas, inconsistencias teóricas, vacíos documentales, patrones estadísticos o resultados que contradicen expectativas anteriores. Los sistemas computacionales permiten combinar estas fuentes y examinar numerosas relaciones entre variables, conceptos o mecanismos. Esta capacidad resulta útil cuando el campo posee una literatura extensa o datos demasiado complejos para una exploración individual. No obstante, la procedencia de la hipótesis debe mantenerse identificable, porque una propuesta derivada de asociaciones estadísticas requiere un tratamiento distinto de aquella sustentada en principios teóricos o evidencia experimental acumulada.

La identificación automatizada de relaciones constituye una fuente frecuente de hipótesis, especialmente cuando los modelos detectan asociaciones entre variables, eventos o entidades científicas. Cornelio et al. (2023) muestran que la combinación de evidencia empírica y conocimiento previo puede orientar la búsqueda de relaciones compatibles con principios establecidos. El investigador debe distinguir si el sistema ha encontrado una regularidad descriptiva, una posible relación causal o una expresión matemática ajustada a datos. Esta clasificación inicial determina qué tipo de evidencia será necesaria para evaluar posteriormente la propuesta.

La generación de alternativas permite evitar que la investigación se limite a una única explicación formulada desde el comienzo. Los modelos pueden producir hipótesis competidoras, señalar variables omitidas y sugerir mecanismos que expliquen el mismo resultado desde perspectivas diferentes. Zhang et al. (2025) sitúan esta capacidad dentro de un proceso científico en el que los grandes modelos de lenguaje apoyan la ideación, el refinamiento y la evaluación preliminar. La diversidad de propuestas es valiosa cuando amplía el razonamiento; se vuelve problemática si genera opciones abundantes sin criterios para discriminar su calidad.

La priorización comienza con la plausibilidad, entendida como la compatibilidad de la hipótesis con el conocimiento científico disponible. Una propuesta plausible no necesita coincidir completamente con teorías previas, pero debe explicar qué supuestos cuestiona y por qué merece ser evaluada. Cory et al. (2024) evidencian que el conocimiento teórico puede utilizarse para restringir espacios de búsqueda y descartar relaciones incompatibles con principios conocidos. Esta función reduce alternativas arbitrarias, aunque no debe impedir investigar anomalías capaces de revelar limitaciones en los marcos conceptuales existentes.

La novedad y la relevancia constituyen criterios diferentes que deben evaluarse conjuntamente. Una hipótesis puede ser original, pero carecer de importancia para el campo; también puede abordar un problema relevante mediante una variación mínima de conocimientos ya establecidos. La inteligencia artificial puede comparar propuestas con literatura disponible para estimar similitudes y antecedentes, pero no determina por sí sola el valor científico o social del problema. La decisión requiere considerar qué conocimiento podría producirse, quiénes se beneficiarían, qué controversia podría aclararse y cómo la respuesta modificaría la comprensión del fenómeno.

La viabilidad se relaciona con los recursos, datos, tiempo, infraestructura y conocimientos necesarios para comprobar una hipótesis. Los sistemas pueden ayudar a estimar procedimientos, identificar fuentes de información y anticipar dificultades, pero estas aproximaciones deben contrastarse con condiciones reales. Una propuesta científicamente interesante puede no ser evaluable en el contexto disponible, mientras que otra menos ambiciosa puede generar evidencia inmediata y útil. Priorizar no implica descartar definitivamente las hipótesis inviables, sino distinguir cuáles pueden investigarse ahora y cuáles requieren nuevos datos, métodos o capacidades institucionales.

La comprobabilidad exige formular la hipótesis de manera que pueda confrontarse con observaciones, mediciones o experimentos. Deben identificarse variables, resultados esperados y condiciones bajo las cuales la propuesta sería rechazada o modificada. Los sistemas multiagente pueden contribuir a este proceso mediante funciones diferenciadas de generación, crítica y clasificación. Ghareeb et al. (2026) plantean una organización distribuida de tareas científicas, mientras Gottweis et al. (2026) muestran que agentes especializados pueden debatir y refinar hipótesis antes de priorizarlas para validaciones posteriores.

El debate entre agentes permite someter cada propuesta a objeciones, comparaciones y reformulaciones sucesivas. Un agente puede producir la hipótesis, otro revisar antecedentes, un tercero identificar debilidades metodológicas y otro evaluar su potencial experimental. Esta interacción genera una selección más estructurada que una respuesta aislada, pero no elimina los riesgos de consenso artificial. Los agentes pueden compartir sesgos, utilizar información incompleta o reforzar argumentos incorrectos. Por ello, sus clasificaciones deben considerarse recomendaciones justificables y no decisiones definitivas sobre qué líneas científicas merecen desarrollarse.

La validación humana y experimental completa el proceso de priorización. El investigador examina coherencia, pertinencia, riesgos, recursos y consecuencias antes de decidir qué propuesta será sometida a prueba. Posteriormente, los resultados empíricos pueden confirmar parcialmente la hipótesis, exigir su modificación o conducir a su rechazo. La generación automatizada acelera la exploración de alternativas, pero el conocimiento científico surge del contraste sistemático y documentado. Una hipótesis computacionalmente convincente permanece abierta mientras no exista evidencia independiente que permita valorar su alcance, reproducibilidad y capacidad explicativa.

Tabla 5

Criterios para la priorización de hipótesis generadas con inteligencia artificial

Criterio de priorización	Pregunta orientadora	Evidencia necesaria	Riesgo asociado
Plausibilidad	¿La propuesta es compatible con conocimientos, teorías y resultados establecidos?	Antecedentes científicos, principios teóricos y evidencia previa	Aceptar relaciones coherentes en apariencia, pero carentes de fundamento

Criterio de priorización	Pregunta orientadora	Evidencia necesaria	Riesgo asociado
Novedad	¿La hipótesis aporta una relación, explicación o perspectiva no desarrollada previamente?	Comparación sistemática con literatura y líneas de investigación existentes	Confundir redacción original con contribución científica real
Relevancia	¿La comprobación de la hipótesis produciría conocimiento significativo para el campo o la sociedad?	Problemas disciplinares, necesidades sociales y debates científicos	Priorizar temas llamativos, pero de escaso valor explicativo o práctico
Viabilidad	¿Existen datos, recursos, métodos, tiempo e infraestructura para evaluarla?	Disponibilidad de información, capacidades técnicas y estimación de recursos	Seleccionar propuestas imposibles de comprobar en las condiciones existentes
Comprobabilidad	¿La hipótesis puede traducirse en variables, observaciones o resultados contrastables?	Diseño preliminar, indicadores y criterios de aceptación o rechazo	Mantener afirmaciones demasiado amplias, ambiguas o infalsables
Coherencia metodológica	¿La estrategia propuesta permite responder de manera válida a la hipótesis?	Correspondencia entre pregunta, diseño, muestra, técnicas y análisis	Utilizar procedimientos técnicamente sofisticados, pero inadecuados
Potencial explicativo	¿La hipótesis ayuda a comprender relaciones, procesos o mecanismos?	Comparación con explicaciones alternativas y conocimiento causal disponible	Reducir el valor científico a capacidad predictiva
Riesgo y responsabilidad	¿La comprobación puede generar daños, conflictos o usos problemáticos?	Evaluación ética, protección de participantes y análisis de consecuencias	Priorizar eficiencia o novedad sin considerar efectos previsibles

Nota. La tabla presenta criterios para evaluar y ordenar hipótesis generadas o refinadas mediante inteligencia artificial. La clasificación computacional constituye un apoyo preliminar; la selección final requiere valoración científica, metodológica, ética y experimental por parte de los investigadores.

La priorización de hipótesis generadas con inteligencia artificial requiere valorar más que su novedad o apariencia de coherencia. Una propuesta científicamente útil debe ser plausible, relevante, viable, comprobable y metodológicamente consistente, además de aportar capacidad explicativa y considerar sus posibles consecuencias. La inteligencia

artificial puede ayudar a comparar alternativas y organizar criterios, pero no determina por sí sola cuál merece desarrollarse. La selección final corresponde a los investigadores, quienes deben justificarla mediante evidencia, análisis crítico, evaluación ética y posibilidades reales de comprobación.

2.3 Modelos generativos como instrumentos de razonamiento científico

Los modelos generativos pueden apoyar el razonamiento científico, pero su capacidad para producir lenguaje coherente no debe confundirse con comprensión demostrada. Estos sistemas generan respuestas mediante relaciones aprendidas durante el entrenamiento y pueden organizar conceptos, formular alternativas o desarrollar argumentos plausibles. Zhang et al. (2025) señalan que los grandes modelos de lenguaje pueden intervenir en diversas etapas de la investigación, siempre que sus aportes sean sometidos a comprobación. Su valor científico depende menos de la fluidez de la respuesta que de la posibilidad de justificarla mediante datos, reglas y evidencia independiente.

La generación textual constituye el nivel más básico de utilización. En esta función, el modelo redacta explicaciones, resume información, reformula preguntas o propone posibles conexiones entre conceptos. Estas capacidades facilitan la exploración inicial, pero no garantizan que el contenido generado represente un razonamiento válido. Un texto puede conservar coherencia gramatical y estructura argumentativa mientras combina premisas incompatibles o conclusiones insuficientemente fundamentadas. Por ello, la calidad lingüística debe evaluarse de manera separada de la calidad científica, especialmente cuando el resultado pretende incorporarse a una hipótesis, interpretación o explicación.

Un nivel posterior corresponde a la organización de argumentos. El modelo puede distinguir premisas, conclusiones, objeciones y evidencias dentro de una discusión, además de ordenar información dispersa en secuencias comprensibles. Behrend y Landers (2025) muestran que los grandes modelos de lenguaje pueden colaborar en la generación y revisión de materiales de investigación, siempre que los investigadores mantengan control sobre sus decisiones. Aplicado al razonamiento, este apoyo permite comparar posiciones y detectar componentes ausentes, pero la solidez del argumento continúa dependiendo de la calidad y pertinencia de las premisas utilizadas.

Los modelos generativos también pueden comparar explicaciones alternativas para un mismo fenómeno. Esta función es útil cuando la investigación requiere valorar distintas interpretaciones, mecanismos o relaciones posibles. El sistema puede exponer coincidencias, contradicciones y consecuencias esperadas de cada propuesta, ayudando al investigador a ampliar el espacio de análisis. Sin embargo, la comparación no debe limitarse a la consistencia interna. Una explicación científicamente útil necesita conectarse con observaciones, teorías y pruebas que permitan distinguirla de otras alternativas, incluso cuando todas parezcan razonables desde el punto de vista discursivo.

El razonamiento científico mejora cuando los modelos operan con información externa verificable, en lugar de depender únicamente de sus parámetros internos. Cornelio et al. (2023) integran datos experimentales y conocimiento teórico para identificar relaciones matemáticas derivables desde principios establecidos. Este enfoque muestra que las reglas científicas pueden restringir el espacio de respuestas y evitar que el sistema valore únicamente el ajuste numérico. La incorporación de axiomas, bases documentales y resultados medidos permite construir procesos más controlados, porque las propuestas deben satisfacer condiciones que trascienden la plausibilidad lingüística.

La conexión con herramientas amplía estas capacidades mediante búsquedas documentales, ejecución de código, cálculos y consulta de bases estructuradas. Boiko et al. (2023) demostraron que un modelo de lenguaje puede coordinar componentes especializados para planificar procedimientos químicos y acceder a recursos técnicos. En este caso, el modelo funciona como una interfaz que distribuye tareas y utiliza los resultados de otras herramientas para continuar el proceso. Su desempeño depende, por tanto, de la fiabilidad de cada componente, de la correcta transferencia de información y de los controles establecidos por los investigadores.

El razonamiento estructurado organiza el problema en pasos, subpreguntas y comprobaciones intermedias. En lugar de solicitar una conclusión inmediata, el investigador puede exigir que el sistema identifique supuestos, reúna evidencia, compare alternativas y revise inconsistencias. Gottweis et al. (2026) aplican una arquitectura multiagente en la que diferentes componentes generan, critican, clasifican y refinan hipótesis. La interacción entre funciones especializadas contribuye a mejorar las propuestas, aunque el intercambio entre agentes no garantiza objetividad, porque varios sistemas pueden compartir limitaciones, sesgos o fuentes de información incompletas.

La explicación científica exige algo más que describir una asociación o reproducir una regularidad. Krenn et al. (2022) sostienen que la inteligencia artificial debe evaluarse por su capacidad de favorecer comprensión científica, lo cual implica relacionar resultados con conceptos, mecanismos y principios. Un modelo puede predecir correctamente sin explicar por qué ocurre un fenómeno, del mismo modo que puede producir una explicación convincente sin demostrarla. La investigación debe distinguir desempeño predictivo, interpretación y explicación, debido a que cada uno responde a preguntas diferentes y requiere criterios específicos de evaluación.

Las limitaciones se vuelven especialmente importantes cuando se pretende extraer relaciones causales. Los modelos generativos pueden proponer mecanismos, pero no establecen causalidad por la sola organización de información previa. Para sostener una relación causal se necesitan supuestos identificables, diseños adecuados, intervenciones o evidencia compatible con explicaciones alternativas. La capacidad del modelo para completar patrones puede favorecer hipótesis causales preliminares, aunque también puede reproducir asociaciones frecuentes sin distinguir su origen. La evaluación disciplinar debe determinar qué afirmaciones son descriptivas, predictivas, explicativas o causales.

En consecuencia, los modelos generativos pueden funcionar como instrumentos de razonamiento cuando se integran en procesos que incluyen información verificable, reglas explícitas, herramientas y validación externa. Su aporte se encuentra en organizar argumentos, ampliar alternativas, coordinar operaciones y formular propuestas susceptibles de contraste. No obstante, el descubrimiento científico exige que las conclusiones generadas sean confrontadas con datos, experimentos y conocimiento especializado. El modelo contribuye al razonamiento, pero la validez del conocimiento depende de procedimientos capaces de demostrar que la respuesta es coherente, comprobable y científicamente relevante.

Figura 6

Niveles de utilización de modelos generativos en el razonamiento científico



Nota. La figura representa niveles progresivos de utilización de modelos generativos. La generación textual constituye una capacidad inicial, mientras que el razonamiento científico requiere integrar información verificable, herramientas especializadas, criterios explícitos y validación externa.

Los modelos generativos pueden contribuir al razonamiento científico en distintos niveles, desde la producción lingüística hasta la organización, comparación y refinamiento de propuestas. Su aporte aumenta cuando integran literatura, datos, cálculos, código y herramientas especializadas dentro de procesos estructurados. Sin embargo, ninguna secuencia de razonamiento generada constituye por sí sola una demostración científica. La validez depende de contrastar las respuestas con evidencia verificable, conocimiento disciplinar, experimentos y revisión humana, manteniendo separados la plausibilidad argumentativa y el conocimiento científicamente comprobado.

2.4 Descubrimiento de patrones en datos complejos

La inteligencia artificial amplía la capacidad científica para descubrir regularidades en conjuntos de datos cuyo volumen, diversidad o número de variables dificultan el análisis convencional. Estos datos pueden combinar mediciones experimentales, imágenes, textos, señales y registros longitudinales, generando relaciones imposibles de reconocer mediante inspección directa. Guerrero et al. (2025) destacan que la IA facilita el

procesamiento de información científica extensa y la identificación de tendencias relevantes. Sin embargo, encontrar un patrón constituye el comienzo de la interpretación, no la demostración automática de su significado o validez.

La complejidad de los datos no depende únicamente de su cantidad. También puede originarse en la alta dimensionalidad, es decir, en la presencia simultánea de numerosas variables relacionadas entre sí. En estas condiciones, los modelos construyen representaciones que condensan la información y permiten localizar estructuras internas. Truhn et al. (2024) muestran que los modelos multimodales pueden integrar datos clínicos, imágenes y componentes moleculares. Estas representaciones facilitan el análisis conjunto, aunque pueden ocultar cómo determinadas variables influyeron en el resultado producido por el sistema.

La clasificación permite asignar observaciones a categorías definidas previamente, como tipos de documentos, estados de un proceso o grupos con características específicas. Para entrenar un clasificador se requieren ejemplos etiquetados cuya calidad condiciona directamente el resultado. Si las categorías son ambiguas, incompletas o reflejan criterios sesgados, el sistema reproducirá esas limitaciones. Muñoz et al. (2025) señalan que la inteligencia artificial fortalece el tratamiento de información y la producción científica, pero su aplicación exige controles sobre los datos empleados y los procedimientos de evaluación.

El agrupamiento cumple una función diferente porque identifica conjuntos de observaciones similares sin partir necesariamente de categorías preestablecidas. Esta capacidad permite explorar estructuras emergentes, segmentar información e identificar perfiles que después pueden examinarse desde una perspectiva disciplinar. No obstante, los grupos obtenidos dependen de las variables seleccionadas, las representaciones construidas y los criterios de similitud aplicados. Un agrupamiento computacional no constituye por sí mismo una categoría científica; debe analizarse si sus componentes comparten propiedades relevantes y si la separación conserva estabilidad en muestras distintas.

La detección de anomalías se orienta a localizar observaciones que se apartan de los comportamientos predominantes. Estos casos pueden representar errores de medición, registros alterados o fenómenos poco frecuentes con interés científico. Rodríguez et al. (2025) identifican el reconocimiento de patrones y la automatización analítica entre las

tendencias de aplicación de la IA en la investigación. La utilidad de una anomalía depende de su revisión contextual, porque eliminar automáticamente los valores atípicos podría borrar evidencia importante, mientras conservar errores no identificados puede distorsionar las conclusiones.

Las asociaciones multivariadas permiten examinar relaciones que surgen de la interacción simultánea entre numerosos factores. Un modelo puede identificar combinaciones predictivas que pasarían inadvertidas al estudiar cada variable de manera aislada. Sin embargo, una asociación estable no demuestra que una variable produzca cambios en otra. Factores no observados, errores de selección o dependencias temporales pueden originar correlaciones aparentemente significativas. Gamella et al. (2025) subrayan la importancia de evaluar metodologías de descubrimiento causal mediante entornos donde puedan realizarse intervenciones controladas y comparaciones con mecanismos conocidos.

Uno de los principales riesgos es el sobreajuste, que ocurre cuando el modelo aprende particularidades del conjunto utilizado durante el entrenamiento y presenta un rendimiento insuficiente ante datos nuevos. El sistema puede parecer preciso porque reproduce patrones específicos de la muestra, aunque carezca de capacidad para generalizar. Barberis et al. (2024) destacan la necesidad de aplicar procedimientos estandarizados que permitan evaluar robustez y reproducibilidad en modelos científicos. La generalización debe comprobarse utilizando datos independientes, diferentes periodos o poblaciones no empleadas durante el desarrollo inicial.

La calidad de los datos determina qué patrones pueden descubrirse y cuáles quedarán ocultos o deformados. Registros incompletos, instrumentos inconsistentes, errores de etiquetado y muestras poco representativas afectan los resultados incluso cuando el modelo posee un alto rendimiento técnico. También deben documentarse las transformaciones realizadas antes del análisis, porque la limpieza, normalización o exclusión de casos modifica la estructura examinada. La evaluación científica necesita reconstruir cómo se obtuvo el conjunto final y determinar si conserva las características necesarias para responder a la pregunta planteada.

La validación externa permite comprobar si el patrón se mantiene en condiciones diferentes de aquellas utilizadas para identificarlo. Puede incluir nuevas muestras, otros equipos de medición, escenarios geográficos distintos o comparaciones con métodos alternativos. Truhn et al. (2024) advierten que los modelos aplicados a información

multimodal deben evaluarse considerando poblaciones y contextos variados. Cuando el resultado cambia significativamente fuera del conjunto original, debe limitarse su alcance. Comunicar esas restricciones evita presentar una regularidad localizada como si fuera una propiedad general del fenómeno científico.

En consecuencia, el descubrimiento de patrones mediante inteligencia artificial requiere diferenciar entre estructura estadística, capacidad predictiva y explicación científica. Los modelos pueden clasificar, agrupar, reconocer anomalías y revelar asociaciones complejas, pero el investigador debe determinar qué relaciones poseen significado disciplinar y cuáles podrían responder al azar, al diseño de los datos o al procedimiento utilizado. El patrón se convierte en conocimiento útil cuando puede reproducirse, interpretarse y contrastarse mediante evidencia independiente, no cuando únicamente alcanza un rendimiento elevado dentro de una muestra determinada.

Tabla 6

Funciones de la inteligencia artificial para el descubrimiento de patrones en datos complejos

Técnica o función	Tipo de patrón identificado	Ejemplo de aplicación científica	Riesgo principal	Método de validación
Clasificación	Correspondencia entre observaciones y categorías previamente definidas	Clasificación de imágenes, documentos o estados de un proceso	Reproducción de errores y sesgos presentes en las etiquetas	Comparación con expertos y evaluación en datos externos
Agrupamiento	Conjuntos de observaciones con características similares	Identificación de perfiles, subtipos o grupos emergentes	Interpretar agrupaciones estadísticas como categorías científicas definitivas	Repetición con diferentes muestras, variables y criterios de similitud
Detección de anomalías	Casos alejados del comportamiento predominante	Localización de fallas, fenómenos raros o mediciones atípicas	Eliminar evidencia relevante o conservar errores de registro	Revisión contextual, repetición de mediciones y contraste con fuentes originales
Reducción de dimensionalidad	Representaciones condensadas de	Visualización y análisis de datos	Pérdida de información e interpretación	Comparación entre métodos y evaluación

Técnica o función	Tipo de patrón identificado	Ejemplo de aplicación científica	Riesgo principal	Método de validación
	múltiples variables	con alta dimensionalidad	limitada de las representaciones	de la información conservada
Integración multimodal	Relaciones entre textos, imágenes, señales y datos estructurados	Combinación de información clínica, visual y molecular	Desigualdad en la calidad de las modalidades y dificultad de trazabilidad	Evaluación separada y conjunta de cada modalidad
Análisis multivariado	Asociaciones producidas por la interacción de numerosos factores	Predicción de resultados complejos mediante múltiples variables	Confusión entre asociación, predicción y causalidad	Validación externa, análisis de sensibilidad y control de variables
Detección de patrones temporales	Tendencias, secuencias y cambios a través del tiempo	Predicción de evolución, fallas o comportamientos futuros	Dependencia excesiva de condiciones históricas	Evaluación prospectiva y comparación en periodos distintos
Descubrimiento causal	Relaciones compatibles con intervenciones y mecanismos	Identificación de factores que producen cambios observables	Inferir causalidad desde correlaciones estadísticas	Experimentos, intervenciones controladas y análisis contrafactual

Nota. La tabla diferencia funciones utilizadas para reconocer estructuras en datos complejos y señala los controles necesarios para evitar que una regularidad estadística sea interpretada automáticamente como explicación científica.

La inteligencia artificial permite reconocer patrones mediante clasificación, agrupamiento, detección de anomalías, reducción de dimensionalidad, análisis temporal e integración de múltiples fuentes. Estas funciones amplían la capacidad para explorar datos complejos y generar hipótesis, pero sus resultados no deben interpretarse automáticamente como categorías científicas, relaciones causales o explicaciones definitivas. La validez depende de comparar métodos, revisar contextos, utilizar datos externos y aplicar controles experimentales cuando corresponda.

2.5 Integración de conocimiento basado en datos y teoría

Los modelos basados exclusivamente en datos pueden identificar regularidades, clasificar observaciones y generar predicciones con alto rendimiento, pero no siempre explican por qué ocurre un fenómeno. Cuando el aprendizaje depende únicamente de asociaciones

empíricas, existe el riesgo de producir resultados difíciles de interpretar o incompatibles con principios científicos conocidos. Krenn et al. (2022) sostienen que la inteligencia artificial adquiere mayor valor para la ciencia cuando contribuye a la comprensión y no solo a la predicción. Esta perspectiva exige relacionar evidencia, conceptos y mecanismos dentro de modelos científicamente justificables.

La información empírica proporciona registros sobre lo observado, mientras que la teoría organiza esos registros mediante conceptos, relaciones y principios. Ambos componentes cumplen funciones distintas y complementarias. Los datos permiten contrastar expectativas y descubrir comportamientos imprevistos; la teoría orienta qué variables observar, cómo interpretar sus relaciones y qué resultados considerar plausibles. Un sistema que integra ambas fuentes puede reducir explicaciones arbitrarias y priorizar aquellas compatibles con el conocimiento disponible, sin impedir que nuevas evidencias revelen insuficiencias, excepciones o contradicciones dentro de los marcos teóricos existentes.

El conocimiento previo puede incorporarse mediante ecuaciones, reglas, axiomas, relaciones causales, límites físicos o restricciones derivadas de la experiencia disciplinar. Estas condiciones actúan como guías durante el aprendizaje y disminuyen el número de soluciones que el modelo debe evaluar. Cornelio et al. (2023) muestran que los sistemas orientados al descubrimiento científico pueden buscar expresiones matemáticas ajustadas a los datos y, al mismo tiempo, derivables desde principios establecidos. De este modo, la precisión empírica se combina con una exigencia de coherencia teórica que fortalece la interpretación.

Las restricciones científicas también ayudan a impedir que un modelo produzca resultados numéricamente adecuados, pero imposibles desde el punto de vista disciplinar. En física pueden expresarse como leyes de conservación; en química, como condiciones de estabilidad o compatibilidad molecular; y en biología, como relaciones funcionales conocidas. Estas restricciones no deben aplicarse de manera rígida, porque una teoría incompleta podría excluir descubrimientos legítimos. Su función consiste en orientar la búsqueda, hacer explícitos los supuestos y permitir que las desviaciones sean evaluadas como posibles errores o anomalías científicamente relevantes.

Los modelos híbridos combinan procedimientos estadísticos, aprendizaje automático y representaciones explícitas del conocimiento. En lugar de elegir entre enfoques basados

en datos o teoría, buscan aprovechar las fortalezas de ambos. Cory et al. (2024) plantean que la evidencia puede complementar principios incompletos y que el conocimiento previo puede compensar conjuntos de datos escasos o ruidosos. Esta relación bidireccional resulta especialmente importante en ciencias donde obtener nuevas observaciones es costoso, lento o técnicamente difícil, pero existen marcos conceptuales capaces de orientar la inferencia.

La integración también permite examinar la coherencia de las propuestas generadas. Una relación puede ajustarse a las observaciones disponibles y, sin embargo, entrar en conflicto con otros resultados o principios bien establecidos. El modelo debe mostrar qué datos respaldan la propuesta, qué supuestos emplea y bajo cuáles condiciones conserva su validez. Esta trazabilidad facilita comparar explicaciones alternativas y distinguir entre una regularidad localizada y una relación con mayor alcance. La coherencia científica no significa ausencia de contradicción, sino capacidad para identificarla, justificarla y someterla a evaluación.

El descubrimiento de ecuaciones constituye una aplicación representativa de esta integración. Los sistemas pueden analizar observaciones y buscar expresiones capaces de describirlas con precisión, pero el resultado adquiere mayor valor cuando presenta una estructura interpretable y compatible con conocimiento disciplinar. Cornelio et al. (2023) evidencian que una ecuación derivable desde axiomas ofrece ventajas frente a una expresión que únicamente reproduce datos. Aun así, su aceptación requiere examinar unidades, parámetros, condiciones iniciales y límites de aplicación antes de considerarla una contribución científica generalizable.

En la experimentación, la teoría puede orientar la selección de combinaciones prometedoras y los datos obtenidos pueden actualizar las predicciones. Szymanski et al. (2023) muestran esta relación al integrar cálculos, conocimiento previo y resultados experimentales en la búsqueda de materiales. Cada prueba aporta evidencia para ajustar el modelo y decidir qué alternativa conviene explorar posteriormente. El proceso no sigue una dirección única desde la teoría hacia los datos, sino un ciclo en el que predicciones, mediciones y conocimientos acumulados se revisan continuamente según los resultados observados.

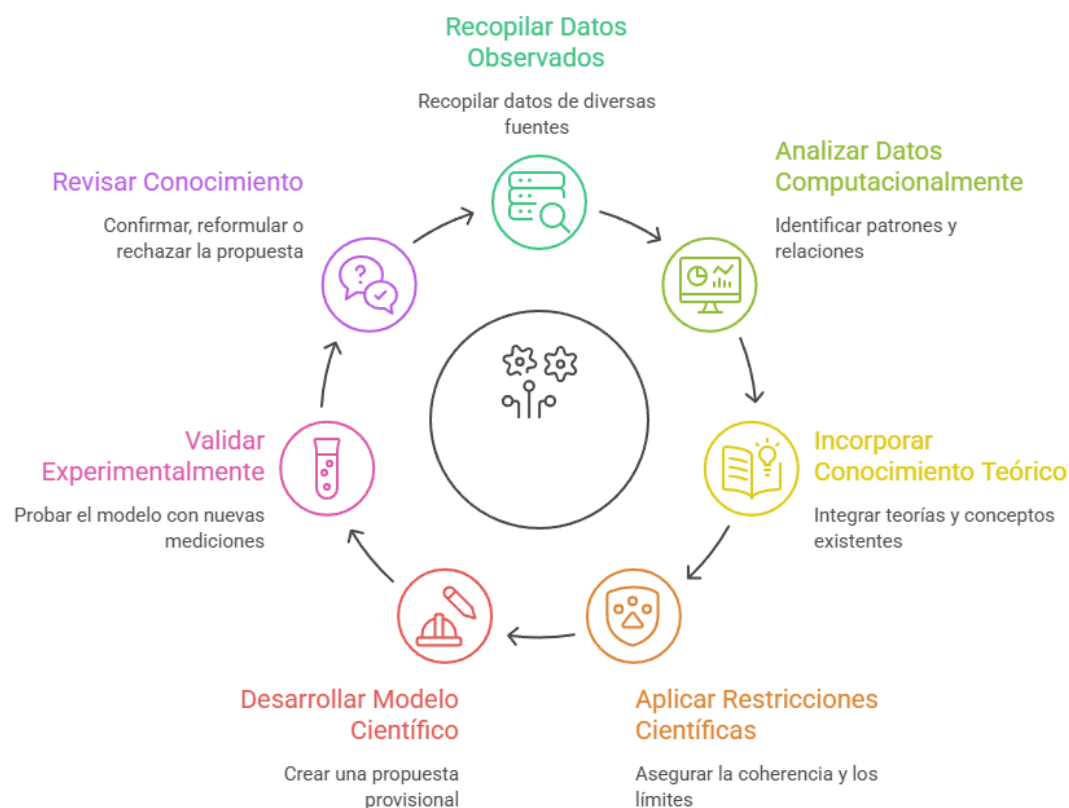
La evidencia experimental conserva una función decisiva porque permite comprobar si las relaciones propuestas se mantienen fuera del entorno computacional. Una simulación

o expresión matemática puede ser coherente con determinados datos y fallar cuando se modifican las condiciones. Por ello, la validación debe incluir nuevas mediciones, replicaciones y comparaciones con explicaciones alternativas. Gamella et al. (2025) resaltan la necesidad de utilizar entornos controlados para evaluar métodos de inferencia y descubrimiento causal. Estos espacios permiten conocer los mecanismos reales y comparar con precisión aquello que el modelo logra recuperar.

La integración entre datos y teoría también favorece el razonamiento causal. Los datos observacionales pueden revelar asociaciones, pero la teoría contribuye a formular mecanismos y anticipar efectos de intervenciones. Cuando ambas fuentes se articulan, es posible distinguir relaciones predictivas de propuestas causales que requieren comprobación adicional. Sin embargo, incorporar conocimiento previo no garantiza que el mecanismo sea correcto. Los supuestos deben permanecer explícitos y ser revisados cuando aparezcan resultados incompatibles.

Figura 7

Integración de datos, teoría y validación en modelos científicos asistidos por inteligencia artificial



Nota. La figura representa la interacción entre evidencia empírica, conocimiento teórico y restricciones científicas. La propuesta generada por el modelo permanece provisional hasta ser contrastada mediante experimentación y revisión disciplinar.

La integración de datos, teoría y validación permite que la inteligencia artificial participe en la construcción de modelos científicos sin reducir el conocimiento a patrones estadísticos. Los datos aportan evidencia, la teoría orienta la interpretación y las restricciones disciplinares delimitan relaciones plausibles. A partir de esta articulación pueden formularse ecuaciones, mecanismos o predicciones provisionales. Sin embargo, su valor científico depende de la validación experimental, la replicación y el contraste independiente, procesos que permiten confirmar, reformular o rechazar las propuestas generadas.

2.6 Simulación, predicción y construcción de escenarios

La inteligencia artificial amplía la capacidad de representar fenómenos, anticipar resultados y examinar condiciones que todavía no han ocurrido. Sin embargo, simulación, predicción, proyección y construcción de escenarios no constituyen procedimientos equivalentes. Cada uno responde a preguntas diferentes, utiliza supuestos específicos y produce resultados con distinto alcance. Suazo (2023) reconoce que la IA puede apoyar la comprobación de hipótesis mediante simulaciones y experimentos virtuales. Estas posibilidades son valiosas siempre que los resultados se interpreten como aproximaciones condicionadas y no como descripciones infalibles del futuro.

La simulación consiste en representar computacionalmente el comportamiento de un sistema bajo determinadas reglas, parámetros y condiciones iniciales. Puede utilizarse para analizar fenómenos físicos, biológicos, sociales o productivos sin intervenir directamente sobre el sistema real. Su utilidad reside en explorar consecuencias, comparar configuraciones y observar dinámicas difíciles de reproducir experimentalmente. No obstante, toda simulación simplifica la realidad. Sus resultados dependen de las relaciones incorporadas, de la calidad de los parámetros y de los elementos que fueron excluidos durante la construcción del modelo.

La predicción busca estimar un resultado desconocido a partir de patrones identificados en datos disponibles. Los modelos pueden anticipar categorías, valores, riesgos o comportamientos futuros cuando las condiciones se mantienen suficientemente próximas a las observadas durante su entrenamiento. Truhn et al. (2024) muestran que la integración de información clínica, visual y molecular puede fortalecer predicciones en contextos de medicina de precisión. Aun así, un modelo predictivo puede alcanzar buen rendimiento

sin explicar los mecanismos que originan el resultado ni conservar su precisión ante poblaciones diferentes.

La proyección extiende tendencias observadas hacia periodos o condiciones futuras, generalmente bajo el supuesto de que determinados comportamientos permanecerán relativamente estables. Puede utilizarse para estimar demanda, crecimiento, propagación o evolución de indicadores. A diferencia de la predicción individual, suele orientarse hacia trayectorias agregadas y horizontes temporales definidos. Su limitación principal aparece cuando se presentan cambios estructurales, innovaciones, crisis o acontecimientos sin precedentes. La prolongación de tendencias históricas no permite anticipar adecuadamente fenómenos que alteran las relaciones sobre las cuales se construyó el modelo.

La construcción de escenarios no pretende establecer un único resultado probable, sino representar varios futuros posibles según combinaciones de supuestos y decisiones. Los escenarios pueden organizarse alrededor de condiciones favorables, intermedias o adversas, aunque su valor aumenta cuando describen mecanismos y variables claramente diferenciados. La inteligencia artificial permite explorar numerosas combinaciones y estimar consecuencias asociadas con cada una. Sin embargo, los escenarios no son predicciones cerradas. Funcionan como instrumentos para examinar incertidumbres, comparar decisiones y reconocer qué factores podrían modificar la evolución del fenómeno estudiado.

Los contrafactuales responden a preguntas sobre lo que habría ocurrido bajo una condición diferente de la observada. Este razonamiento compara el resultado real con una situación hipotética en la que se modifica una intervención, exposición o decisión. Su aplicación resulta relevante para evaluar políticas, tratamientos y mecanismos causales. Gamella et al. (2025) destacan la utilidad de entornos físicos controlados para comprobar metodologías de descubrimiento causal. El principal desafío es que el resultado contrafactual no puede observarse directamente en la misma unidad y debe estimarse mediante diseños, supuestos y comparaciones apropiados.

La intervención se distingue de la simple observación porque implica modificar deliberadamente una variable para examinar sus efectos. Mientras una predicción estima qué podría suceder, el análisis de intervención pregunta qué ocurriría si se realizara una acción concreta. Esta diferencia resulta fundamental para no confundir asociaciones con

relaciones causales. Un modelo puede anticipar correctamente un resultado utilizando variables correlacionadas que no producen el fenómeno. La evaluación causal exige justificar por qué la intervención debería generar un cambio, controlar explicaciones alternativas y comprobar que el efecto se mantiene bajo condiciones comparables.

Cory et al. (2024) muestran que la combinación de datos y conocimiento teórico puede restringir soluciones y favorecer relaciones científicamente coherentes. Esta integración también fortalece simulaciones y escenarios, porque permite incorporar leyes, límites y mecanismos conocidos. Sin embargo, la presencia de teoría no elimina la incertidumbre. Los parámetros pueden estar mal estimados, los axiomas ser incompletos y las condiciones reales diferir de aquellas representadas. Por ello, los modelos deben documentar sus supuestos y distinguir claramente entre información observada, relaciones incorporadas y resultados generados computacionalmente.

El análisis de sensibilidad permite examinar cuánto cambian los resultados cuando se modifican parámetros, datos o condiciones iniciales. Si pequeñas variaciones producen conclusiones completamente distintas, el modelo presenta inestabilidad y sus resultados deben comunicarse con prudencia. Krenn et al. (2022) sostienen que la contribución científica de la inteligencia artificial no puede limitarse a su rendimiento predictivo, sino que debe favorecer comprensión. En este sentido, analizar sensibilidad ayuda a identificar qué variables influyen realmente, cuáles supuestos dominan el resultado y dónde se concentra la incertidumbre del modelo.

La validación exige comparar simulaciones, predicciones y escenarios con datos independientes, observaciones posteriores o resultados experimentales. También debe evaluarse el comportamiento del modelo ante condiciones distintas de las utilizadas durante su construcción. Los eventos no observados representan una limitación relevante, porque los sistemas aprenden principalmente de regularidades disponibles. En consecuencia, estos instrumentos deben apoyar decisiones sin presentarse como certezas. Su valor científico reside en hacer explícitos los supuestos, explorar alternativas, cuantificar incertidumbres y facilitar contrastes que puedan revisarse cuando aparezca nueva evidencia.

Tabla 7

Diferencias entre simulación, predicción y construcción de escenarios asistidas por inteligencia artificial

Concepto	Propósito	Entrada principal	Resultado esperado	Ejemplo científico	Limitación principal
Simulación	Representar el comportamiento de un sistema bajo reglas definidas	Modelos, parámetros, condiciones iniciales y relaciones conocidas	Evolución computacional del sistema	Simular el comportamiento de un material bajo diferentes temperaturas	Depende de simplificaciones y supuestos incorporados
Predicción	Estimar un resultado desconocido a partir de patrones previos	Datos históricos, variables observadas y modelo entrenado	Valor, categoría, riesgo o probabilidad	Estimar la respuesta de un paciente a un tratamiento	Puede predecir sin explicar el mecanismo
Proyección	Extender tendencias hacia un periodo futuro	Serie históricas, tasas de cambio y supuestos de continuidad	Trayectoria esperada de un indicador	Proyectar el crecimiento de una línea de investigación	Falla ante cambios estructurales o eventos inéditos
Escenario	Comparar futuros posibles bajo diferentes condiciones	Supuestos alternativos, variables críticas y decisiones posibles	Conjunto de trayectorias plausibles	Evaluar escenarios de adopción de IA en instituciones científicas	No determina cuál futuro ocurrirá
Contrafactual	Estimar qué habría sucedido bajo una condición diferente	Datos observados, comparación válida y modelo causal	Resultado hipotético alternativo	Estimar qué habría ocurrido sin una intervención	El resultado alternativo no puede observarse directamente
Intervención	Evaluar el efecto de modificar deliberadamente una variable	Diseño experimental o causal, tratamiento y grupo de comparación	Cambio atribuible a una acción	Medir el efecto de alterar una condición experimental	Requiere controlar explicaciones alternativas
Explicación causal	Identificar por qué y mediante qué mecanismo	Teoría, datos, intervenciones y	Mecanismo o relación causal sustentada	Determinar cómo una variable produce	No puede inferirse únicamente desde correlaciones

Concepto	Propósito	Entrada principal	Resultado esperado	Ejemplo científico	Limitación principal
	ocurre un fenómeno	relaciones causales		cambios en otra	

Nota. La tabla diferencia procedimientos que suelen utilizarse conjuntamente, pero responden a propósitos científicos distintos. Las simulaciones, predicciones y escenarios apoyan la exploración, mientras que las afirmaciones causales requieren supuestos explícitos, intervenciones o evidencia adicional.

La simulación, la predicción y la construcción de escenarios cumplen funciones complementarias, pero no equivalentes, dentro de la investigación asistida por inteligencia artificial. Mientras la simulación representa el comportamiento de un sistema y la predicción estima resultados probables, los escenarios comparan futuros posibles bajo condiciones alternativas. Las proyecciones, los contrafactuales y las intervenciones añaden niveles distintos de análisis. Por ello, las afirmaciones causales requieren evidencia más sólida que una tendencia o una predicción, incluyendo supuestos explícitos, diseños adecuados y comprobación independiente.

2.7 Producción de conocimiento en investigación cualitativa

La investigación cualitativa busca comprender significados, experiencias, prácticas y relaciones situadas dentro de contextos sociales específicos. La inteligencia artificial puede apoyar este proceso al organizar grandes volúmenes de entrevistas, documentos, imágenes o registros, pero no reemplaza la interpretación construida por el investigador. Goyanes y Lopezosa (2024) identifican usos de modelos generativos en entrevistas, revisiones y análisis de contenido, aunque advierten que sus resultados continúan siendo preliminares. El aporte tecnológico debe integrarse sin reducir el conocimiento cualitativo a frecuencias, coincidencias lingüísticas o clasificaciones descontextualizadas.

La transcripción constituye una de las primeras tareas susceptibles de asistencia automatizada. Los sistemas pueden convertir grabaciones en texto, separar participantes, reconocer segmentos y generar marcas temporales que faciliten la revisión. Esta automatización reduce trabajo operativo, pero puede introducir errores asociados con acentos, interrupciones, vocabulario especializado o calidad del audio. Por ello, la transcripción no debe considerarse un registro definitivo. El investigador necesita contrastarla con la grabación original, corregir expresiones y conservar elementos como silencios, énfasis o cambios de tono cuando resulten relevantes para el análisis.

La codificación inicial permite identificar fragmentos relacionados con ideas, acciones, emociones o acontecimientos presentes en el corpus. Los modelos generativos pueden

sugerir códigos, localizar pasajes semejantes y aplicar etiquetas preliminares a conjuntos extensos. Perkins y Roe (2024) muestran que ChatGPT puede participar en análisis temáticos inductivos mediante procedimientos guiados. Sin embargo, la codificación automatizada depende de las instrucciones proporcionadas y de la forma en que el sistema interpreta el lenguaje. Los códigos deben revisarse para evitar simplificaciones, asociaciones superficiales o categorías alejadas del propósito investigativo.

La construcción de categorías requiere integrar códigos, reconocer relaciones y establecer niveles de abstracción que expliquen componentes relevantes del fenómeno. La inteligencia artificial puede agrupar etiquetas, proponer jerarquías y mostrar coincidencias entre testimonios, aunque no determina por sí sola qué categorías poseen valor interpretativo. Nguyen (2025) plantea que estos sistemas pueden funcionar como asistentes en el análisis temático cuando el investigador conserva la dirección conceptual. Las categorías deben fundamentarse en los datos, el marco teórico y el contexto, en lugar de aceptarse únicamente porque aparecen con frecuencia o presentan semejanza semántica.

El análisis temático supone identificar patrones de significado que atraviesan el material y responden a las preguntas de investigación. Bennis y Mouwafaq (2025) examinan el uso de inteligencia artificial para procesar conjuntos cualitativos amplios y apoyar la generación de temas. Esta capacidad puede facilitar comparaciones y revelar conexiones poco visibles, pero cada tema necesita coherencia interna, diferenciación respecto de otros y evidencia suficiente. El investigador debe comprobar si representa realmente la experiencia de los participantes o si surge de una generalización lingüística que elimina contradicciones, matices y casos minoritarios.

Las matrices analíticas permiten comparar participantes, categorías, momentos o fuentes mediante una estructura organizada. La inteligencia artificial puede ayudar a completar tablas, resumir segmentos y localizar semejanzas o diferencias entre casos. Esta función resulta útil cuando el corpus es extenso, siempre que los resúmenes permanezcan vinculados con sus fragmentos originales. Morantes (2023) destaca experiencias educativas donde la IA apoya la sistematización de información investigativa. No obstante, una matriz automatizada debe conservar referencias suficientes para reconstruir cómo se obtuvo cada síntesis y revisar posibles pérdidas de significado.

La comparación de testimonios ayuda a reconocer regularidades, contrastes, excepciones y trayectorias particulares. Los modelos pueden señalar expresiones recurrentes o identificar posiciones opuestas, pero sus resultados suelen privilegiar similitudes lingüísticas. Dos participantes pueden utilizar palabras diferentes para describir experiencias equivalentes, o emplear términos semejantes con sentidos distintos. Por ello, comparar no consiste únicamente en medir proximidad textual. El investigador debe considerar historia, ubicación social, relaciones, circunstancias y propósitos comunicativos para comprender por qué los discursos coinciden, divergen o cambian durante el proceso estudiado.

La interpretación transforma códigos y categorías en explicaciones sobre el fenómeno. Esta operación exige relacionar datos, teoría, contexto y posicionamiento del investigador. Girdwood y Manikonda (2026) sostienen que la incorporación de IA en investigación cualitativa requiere decisiones éticas y prácticas desde la conceptualización hasta la publicación. Un sistema puede generar lecturas alternativas o cuestionamientos útiles, pero no posee acceso directo a la experiencia vivida ni a todas las condiciones del campo. Las interpretaciones propuestas deben entenderse como insumos para deliberación y no como conclusiones autónomas.

La reflexividad permite examinar cómo las decisiones, experiencias y supuestos del investigador influyen en la producción del conocimiento. La inteligencia artificial no elimina esta influencia, sino que añade nuevas decisiones relacionadas con selección de herramientas, instrucciones, versiones y criterios de aceptación. Goyanes y Lopezosa (2024) subrayan la importancia del pensamiento crítico en la colaboración humano-máquina. El investigador debe registrar cuándo utilizó IA, qué materiales proporcionó, qué resultados descartó y cómo esa interacción afectó la construcción de códigos, categorías, temas y explicaciones.

La triangulación fortalece la interpretación mediante la comparación de fuentes, investigadores, métodos o perspectivas teóricas. Los resultados generados con inteligencia artificial pueden contrastarse con codificación manual, revisión por pares, entrevistas de validación o análisis de casos discrepantes. Perkins y Roe (2024) muestran que la utilidad del apoyo generativo aumenta cuando se integra en procedimientos transparentes y revisables. La coincidencia entre sistema e investigador no demuestra automáticamente validez; puede reflejar supuestos compartidos. También deben

examinarse desacuerdos, porque pueden revelar ambigüedades, sesgos o dimensiones no consideradas inicialmente.

En consecuencia, la inteligencia artificial puede mejorar la organización, exploración y comparación de materiales cualitativos, pero la producción de conocimiento continúa dependiendo de decisiones interpretativas justificadas. La supervisión humana debe conservar el contexto, reconocer voces minoritarias y valorar contradicciones que un sistema podría tratar como ruido. La calidad del análisis no se mide por la cantidad de texto procesado, sino por la coherencia entre preguntas, datos, categorías y explicaciones. La asistencia automatizada resulta valiosa cuando amplía la reflexión sin sustituir la responsabilidad intelectual del equipo investigador.

Tabla 8

Intervención de la inteligencia artificial en las fases de la investigación cualitativa

Fase cualitativa	Aporte posible de la inteligencia artificial	Decisión humana indispensable	Riesgo interpretativo
Transcripción	Conversión de audio en texto, segmentación e identificación preliminar de hablantes	Corregir errores y decidir qué elementos paralingüísticos conservar	Pérdida de silencios, énfasis, acentos y expresiones contextuales
Organización del corpus	Clasificación de documentos, entrevistas y registros	Definir unidades de análisis y criterios de agrupación	Organizar materiales según semejanzas superficiales
Codificación inicial	Sugerencia y aplicación preliminar de etiquetas	Revisar pertinencia, significado y relación con las preguntas	Imposición de códigos genéricos o alejados del contexto
Construcción de categorías	Agrupamiento de códigos y propuesta de jerarquías	Establecer coherencia conceptual y nivel de abstracción	Confundir frecuencia o similitud semántica con relevancia
Análisis temático	Identificación de patrones y relaciones entre segmentos	Validar temas mediante evidencia suficiente y casos discrepantes	Homogeneización de experiencias y eliminación de contradicciones
Elaboración de matrices	Síntesis y comparación de casos, fuentes o momentos	Mantener vínculo con datos originales y revisar resúmenes	Pérdida de matices durante la condensación de información
Comparación de testimonios	Localización de similitudes, diferencias y recurrencias	Interpretar discursos según circunstancias sociales e históricas	Equiparar semejanza lingüística con equivalencia de significado

Fase cualitativa	Aporte posible de la Decisión inteligencia artificial indispensable	humana	Riesgo interpretativo
Interpretación	Generación de lecturas alternativas relaciones conceptuales	Construir y explicaciones sustentadas en datos, teoría y contexto	Aceptar interpretaciones plausibles sin fundamento suficiente
Reflexividad	Registro de interacciones, decisiones y versiones utilizadas	Examinar cómo la IA modificó el proceso analítico	Ocultar la influencia tecnológica sobre categorías y conclusiones
Triangulación	Comparación entre coincidencias, análisis automatizado, manual y otras fuentes	Valorar coincidencias, desacuerdos y evidencia contradictoria	Tomar el consenso humano-máquina como prueba automática de validez

Nota. La tabla muestra que la inteligencia artificial puede apoyar tareas de organización y análisis, pero las decisiones sobre significado, contexto, validez y construcción interpretativa permanecen bajo responsabilidad humana.

La inteligencia artificial puede apoyar la investigación cualitativa mediante la transcripción, organización, codificación, comparación y síntesis de grandes volúmenes de información. Estas funciones agilizan el trabajo analítico, pero no determinan por sí solas el significado de los testimonios ni la validez de las categorías construidas. La interpretación requiere conservar el vínculo con los datos originales, atender al contexto y examinar casos discrepantes. Por ello, la reflexividad, la triangulación y la revisión humana siguen siendo indispensables para evitar simplificaciones y mantener la coherencia del análisis.

Capítulo

3

Automatización de los procesos de investigación científica

Verónica Luisana López Mendizábal



3.1 Automatización de revisiones sistemáticas

Las revisiones sistemáticas exigen procedimientos transparentes para localizar, seleccionar, evaluar y sintetizar estudios relacionados con una pregunta claramente formulada. La inteligencia artificial puede reducir la carga operativa de varias etapas, pero no modifica la necesidad de aplicar criterios metodológicos explícitos. Li et al. (2024) muestran que los grandes modelos de lenguaje pueden apoyar el cribado de resúmenes mediante criterios definidos previamente. Su utilidad depende de mantener registros verificables y de conservar la intervención humana en aquellas decisiones capaces de modificar la composición final de la evidencia analizada.

El proceso comienza con una pregunta suficientemente delimitada para definir población, fenómeno, intervención, comparación, contexto o resultados, según el tipo de revisión. La IA puede ayudar a identificar conceptos relacionados y convertirlos en términos de búsqueda, pero no debe establecer por sí sola el alcance de la investigación. La pregunta determina qué estudios serán pertinentes y qué información deberá extraerse posteriormente. Una delimitación imprecisa produce búsquedas extensas, criterios ambiguos y decisiones difíciles de reproducir, aunque las etapas posteriores se ejecuten mediante sistemas técnicamente avanzados.

La búsqueda bibliográfica puede automatizar parcialmente la combinación de términos, consulta de bases y recuperación de registros. Goyanes y Lopezosa (2024) reconocen el valor de estas tecnologías para apoyar revisiones y organizar literatura científica. Sin embargo, las estrategias deben adaptarse a cada fuente, porque las bases utilizan campos, operadores y vocabularios diferentes. El investigador necesita revisar si la consulta representa adecuadamente la pregunta y si incluye sinónimos, variantes lingüísticas y términos disciplinares. La automatización agiliza la recuperación, pero no garantiza exhaustividad ni equilibrio documental.

Después de reunir los registros, la deduplicación permite identificar versiones repetidas de una misma publicación. Los sistemas pueden comparar títulos, autores, identificadores y metadatos para eliminar coincidencias exactas o aproximadas. Esta tarea parece sencilla, pero pueden existir preprints, artículos finales, traducciones o registros con información incompleta. Una eliminación incorrecta puede descartar versiones relevantes o conservar duplicados que alteren los conteos. Por ello, los criterios utilizados deben

documentarse y los casos dudosos deben revisarse antes de consolidar la base definitiva para el cribado.

El cribado de títulos y resúmenes constituye una de las etapas con mayor potencial de semiautomatización. Syriani et al. (2024) evaluaron el empleo de ChatGPT para clasificar artículos conforme a criterios de inclusión y exclusión. Los modelos pueden priorizar registros y reducir el número de documentos que requieren revisión inmediata. No obstante, su desempeño depende de la claridad de los criterios, la calidad de los resúmenes y la formulación de las instrucciones. Los estudios inciertos deben mantenerse para una evaluación humana más detallada, evitando exclusiones irreversibles.

La evaluación de elegibilidad mediante texto completo exige determinar si cada estudio cumple todas las condiciones definidas en el protocolo. En esta fase, la inteligencia artificial puede localizar información sobre diseño, muestra, intervención, variables o resultados, pero la decisión final requiere interpretar el documento completo. Li et al. (2024) advierten que el rendimiento de los modelos varía entre tareas y configuraciones. Una publicación puede parecer pertinente según su resumen y quedar excluida por características metodológicas que solo aparecen en métodos, anexos o resultados.

La extracción de datos consiste en registrar información homogénea sobre cada estudio incluido. Los sistemas pueden identificar autores, año, diseño, participantes, variables, instrumentos y resultados, además de trasladarlos a matrices estructuradas. Esta automatización debe acompañarse de reglas para resolver datos ausentes, múltiples mediciones o formatos incompatibles. Las cifras, unidades y descripciones metodológicas necesitan comprobarse directamente en el texto. Un error de extracción puede propagarse hacia la síntesis y modificar las conclusiones, incluso cuando la selección de estudios haya sido realizada correctamente.

La clasificación posterior organiza los estudios por diseños, poblaciones, intervenciones, enfoques o tipos de resultado. La inteligencia artificial puede sugerir agrupaciones y resumir características comunes, facilitando la preparación de una síntesis narrativa. Sin embargo, la similitud lingüística no siempre implica equivalencia metodológica. Dos investigaciones pueden emplear términos semejantes y medir fenómenos diferentes. El equipo revisor debe establecer categorías conceptuales y verificar que los estudios agrupados sean comparables. La automatización funciona como apoyo organizativo, no como sustituto del razonamiento necesario para interpretar heterogeneidad.

La síntesis integra los resultados y determina qué conclusiones pueden sostenerse a partir del conjunto de evidencia. Los modelos generativos pueden producir resúmenes preliminares, pero deben conservar diferencias entre estudios, incertidumbres y resultados contradictorios. También es indispensable registrar las razones de exclusión, porque permiten comprender cómo se obtuvo la muestra documental final. Goyanes y Lopezosa (2024) subrayan que el uso responsable de IA requiere supervisión crítica. Una revisión transparente debe permitir reconstruir las decisiones tomadas y distinguir claramente entre información extraída y elaboración interpretativa.

Figura 8

Flujo de automatización asistida en una revisión sistemática



Nota. La figura representa un flujo de revisión sistemática semiautomatizada. La inteligencia artificial puede apoyar la recuperación, clasificación y extracción de información, mientras que las decisiones de alcance, elegibilidad, interpretación y síntesis permanecen bajo responsabilidad del equipo investigador.

La asistencia automatizada ejecuta tareas específicas; la semiautomatización combina recomendaciones del sistema con decisiones humanas; y la automatización total delega casi todo el proceso. En revisiones sistemáticas, el segundo modelo resulta actualmente más defendible porque equilibra eficiencia y control metodológico. La IA puede priorizar, clasificar y extraer información, mientras los investigadores verifican criterios, resuelven discrepancias y validan la síntesis. La calidad no depende de automatizar más etapas, sino

de documentar cada intervención y conservar responsabilidad humana sobre la evidencia finalmente incluida.

La automatización asistida puede agilizar la búsqueda, deduplicación, clasificación y extracción de información en una revisión sistemática, pero no elimina las decisiones metodológicas que determinan su calidad. La formulación de la pregunta, la aplicación de criterios de elegibilidad, la interpretación de resultados y el registro de exclusiones requieren supervisión humana documentada. Este modelo semiautomatizado permite reducir cargas operativas sin perder trazabilidad, siempre que cada salida sea verificada directamente en las fuentes y las decisiones puedan reconstruirse con claridad.

3.2 Recolección y preparación automatizada de datos

La recolección y preparación de datos constituyen etapas decisivas porque determinan la calidad de los análisis posteriores. La inteligencia artificial puede intervenir desde la creación de instrumentos hasta la integración de bases heterogéneas, reduciendo tareas repetitivas y facilitando el tratamiento de grandes volúmenes de información. Cisneros et al. (2022) destacan que las técnicas e instrumentos deben seleccionarse según los objetivos, el enfoque y las unidades de análisis. La automatización no corrige un diseño deficiente; únicamente procesa con mayor rapidez las decisiones metodológicas adoptadas previamente.

En la generación de instrumentos, los modelos pueden proponer preguntas, escalas, categorías de observación o estructuras preliminares para formularios. Behrend y Landers (2025) examinan el uso de grandes modelos de lenguaje para producir materiales destinados a encuestas y experimentos. Este apoyo puede ampliar alternativas y mejorar la redacción inicial, pero cada ítem debe revisarse según su claridad, pertinencia, nivel de medición y correspondencia con las variables estudiadas. Un instrumento lingüísticamente correcto puede medir de manera inadecuada el constructo que pretende representar.

La captura automatizada comprende la recepción de respuestas digitales, registros de sensores, documentos, imágenes y otros datos producidos durante la investigación. Los sistemas pueden extraer información desde diferentes formatos y convertirla en estructuras analizables. Esta capacidad reduce la transcripción manual, aunque exige comprobar que los campos, unidades y marcas temporales se hayan conservado

correctamente. Cuando los datos proceden de fuentes diversas, también debe documentarse su origen, fecha, forma de obtención y cualquier modificación realizada antes de incorporarlos al conjunto principal.

La digitalización permite transformar materiales impresos, audios o registros no estructurados en datos procesables. La inteligencia artificial puede reconocer texto, segmentar documentos y organizar contenidos, pero su rendimiento varía según la calidad del archivo, el idioma y la complejidad visual. Los errores pueden afectar nombres, cifras, símbolos o categorías relevantes sin ser evidentes durante una revisión superficial. Por ello, las conversiones deben comprobarse mediante muestras, comparación con los originales y reglas de control capaces de detectar valores improbables o inconsistencias entre campos relacionados.

La limpieza de datos busca corregir duplicados, formatos incompatibles, errores de registro y observaciones incompletas. Los procedimientos automatizados pueden identificar anomalías, sugerir correcciones y aplicar reglas uniformes. Sin embargo, una diferencia no siempre representa un error. Un valor extremo puede corresponder a un caso científicamente relevante, y una respuesta aparentemente inconsistente puede reflejar condiciones reales del participante o fenómeno. La decisión de corregir, conservar o excluir información requiere revisar su contexto y registrar el criterio empleado para evitar transformaciones silenciosas del conjunto analizado.

El tratamiento de valores ausentes exige determinar por qué falta la información y qué efecto puede producir su ausencia. Los sistemas pueden detectar patrones de omisión, completar datos mediante estimaciones o recomendar estrategias de imputación. No obstante, reemplazar un valor modifica el conjunto original y puede reducir artificialmente la incertidumbre. El investigador debe distinguir entre omisiones aleatorias, fallas instrumentales y ausencias relacionadas con el fenómeno. Toda imputación necesita justificarse, identificarse y compararse con análisis alternativos para valorar si cambia de manera significativa los resultados obtenidos.

La normalización, el etiquetado y la estructuración permiten que fuentes diferentes puedan analizarse conjuntamente. Truhn et al. (2024) muestran la importancia de armonizar información clínica, visual y molecular antes de integrarla en modelos multimodales. La normalización ajusta escalas y formatos, mientras el etiquetado asigna categorías o atributos a las observaciones. Estas operaciones pueden facilitar la

comparación, pero también eliminar diferencias relevantes. Las reglas deben conservar el significado disciplinar de las variables y evitar que la compatibilidad técnica se obtenga mediante una simplificación excesiva de los datos.

La anonimización busca reducir la posibilidad de identificar a personas o instituciones mediante la eliminación, sustitución o generalización de datos sensibles. Este proceso puede automatizarse parcialmente para localizar nombres, direcciones, identificadores y referencias contextuales. Sin embargo, la eliminación de campos directos no siempre evita la reidentificación cuando varias variables se combinan. La protección debe considerar el tipo de información, los riesgos de vinculación y las condiciones de acceso. También debe evaluarse si la anonimización modifica características necesarias para interpretar adecuadamente los resultados.

La integración y documentación completan la preparación al reunir fuentes, describir transformaciones y conservar un registro reproducible del proceso. Shao et al. (2025) muestran que los sistemas de colaboración humano-IA pueden organizar flujos analíticos y favorecer la repetición de procedimientos. Para ello, cada cambio debe quedar asociado con reglas, versiones y responsables. Los errores silenciosos aparecen cuando una transformación automatizada produce resultados plausibles, pero altera unidades, categorías o significados. La eficiencia solo es útil cuando el equipo puede reconstruir qué ocurrió desde la captura inicial hasta la base final.

Tabla 9
Tareas de recolección y preparación automatizada de datos

Tarea de preparación	Automatización posible	Validación requerida	Riesgo principal
Generación de instrumentos	Propuesta de preguntas, escalas, categorías y formatos	Revisión de validez, claridad y correspondencia con variables	Instrumentos bien redactados, pero metodológicamente inadecuados
Captura de datos	Recepción automática de formularios, sensores y registros digitales	Comprobación de campos, unidades, fechas y procedencia	Pérdida o alteración de información durante la transferencia
Digitalización	Conversión de textos, audios e imágenes en datos estructurados	Comparación con originales y revisión de muestras	Errores de reconocimiento difíciles de detectar

Tarea de preparación	Automatización posible	Validación requerida	Riesgo principal
Limpieza	Identificación de duplicados, inconsistencias y formatos incorrectos	Revisión contextual de casos dudosos	Eliminación de valores válidos considerados anómalos
Valores ausentes	Detección de patrones e imputación preliminar	Justificación del método y análisis de sensibilidad	Reducción artificial de incertidumbre o introducción de sesgos
Normalización	Conversión de escalas, unidades y formatos	Confirmación de equivalencias y conservación del significado	Homogeneización de diferencias científicamente relevantes
Etiquetado	Asignación automática o asistida de categorías	Revisión de criterios y consistencia entre evaluadores	Reproducción de errores o sesgos en las etiquetas
Anonimización	Detección y sustitución de identificadores directos	Evaluación del riesgo de reidentificación	Exposición de información mediante combinación de variables
Integración	Unión de bases, modalidades y fuentes documentales	Verificación de correspondencias, claves y compatibilidad	Mezcla de registros no equivalentes o desalineados
Documentación	Registro de reglas, versiones, transformaciones y responsables	Revisión de trazabilidad y posibilidad de reproducción	Procesos opacos que impiden reconstruir la base final

Nota. La tabla sintetiza tareas susceptibles de automatización durante la recolección y preparación de datos. Cada procedimiento requiere controles humanos para conservar la calidad, el significado y la trazabilidad de la información.

La automatización puede agilizar la creación de instrumentos, la captura, digitalización, limpieza, normalización e integración de datos, pero cada tarea requiere controles específicos para evitar pérdidas, alteraciones o interpretaciones erróneas. La calidad de la base final depende de comprobar equivalencias, conservar el significado de las variables, proteger la identidad de los participantes y documentar todas las transformaciones realizadas. La inteligencia artificial aporta eficiencia operativa, mientras la validación humana garantiza que los datos permanezcan metodológicamente adecuados, trazables y reproducibles.

3.3 Análisis estadístico y programación asistida

La inteligencia artificial puede apoyar el análisis estadístico y la programación mediante la generación de código, la organización de procedimientos y la explicación preliminar de resultados. Su incorporación permite que investigadores con distintos niveles técnicos exploren alternativas analíticas y automaticen tareas repetitivas. Ruiz (2024) señala que estas tecnologías pueden intervenir en diversas fases metodológicas, incluido el procesamiento de datos. Sin embargo, la facilidad para producir comandos o resultados no garantiza que el análisis seleccionado corresponda con la pregunta, el diseño y las características de la información disponible.

La generación asistida de código permite traducir instrucciones expresadas en lenguaje natural hacia secuencias ejecutables en entornos estadísticos y lenguajes de programación. El investigador puede solicitar operaciones de limpieza, transformación, modelado o visualización, además de explicaciones sobre cada bloque producido. Shao et al. (2025) muestran que los sistemas colaborativos pueden agilizar flujos analíticos y facilitar el desarrollo de prototipos computacionales. Aun así, el código debe leerse, probarse y documentarse, porque puede contener funciones inadecuadas, dependencias inexistentes o decisiones que alteren silenciosamente los datos.

La selección preliminar de pruebas estadísticas constituye otra aplicación frecuente. Un sistema puede comparar el tipo de variables, el número de grupos, la estructura de las mediciones y el propósito del estudio para sugerir procedimientos posibles. Guerrero et al. (2025) destacan la capacidad de la inteligencia artificial para apoyar el procesamiento y análisis de información científica. No obstante, la recomendación necesita contrastarse con los supuestos de cada prueba, el diseño muestral y la estrategia de inferencia, evitando elegir técnicas únicamente porque resultan accesibles o generan resultados significativos.

En el análisis descriptivo, la programación asistida facilita calcular medidas de tendencia central, dispersión, frecuencias y distribuciones, además de organizar resultados en tablas. Estas operaciones permiten conocer la estructura inicial de los datos y detectar inconsistencias antes de construir modelos más complejos. El sistema también puede proponer interpretaciones preliminares, pero el investigador debe comprobar que las unidades, categorías y denominadores sean correctos. Un resumen numérico puede resultar técnicamente válido y, al mismo tiempo, ofrecer una representación inadecuada cuando la distribución es asimétrica o contiene subgrupos relevantes.

La construcción de modelos requiere decisiones sobre variables, interacciones, transformaciones y criterios de ajuste. Behrend y Landers (2025) explican que los grandes modelos de lenguaje pueden colaborar en la creación de materiales y procedimientos investigativos, siempre bajo supervisión. En estadística, esta colaboración puede incluir la preparación de regresiones, clasificaciones o análisis predictivos. Sin embargo, aumentar la complejidad no mejora automáticamente la explicación. Cada modelo debe justificarse por su relación con la hipótesis, la naturaleza de los datos y el objetivo científico que orienta el estudio.

La ejecución de experimentos computacionales permite comparar configuraciones, algoritmos y parámetros de manera ordenada. Lu et al. (2026) presentan un sistema orientado a automatizar fases amplias de la investigación en inteligencia artificial, desde la generación de ideas hasta la realización de experimentos y el análisis de resultados. Este tipo de automatización incrementa la velocidad de exploración, pero también puede multiplicar errores cuando una decisión inicial es incorrecta. Por ello, cada ejecución debe conservar información sobre datos, código, configuración, resultados y condiciones del entorno computacional utilizado.

La depuración asistida ayuda a localizar errores de sintaxis, variables mal definidas, incompatibilidades y resultados inesperados. Los modelos pueden explicar mensajes de error y sugerir modificaciones, reduciendo el tiempo dedicado a problemas operativos. Sin embargo, solucionar una falla de ejecución no demuestra que el procedimiento sea metodológicamente correcto. Un programa puede funcionar sin interrupciones y aplicar una transformación equivocada, utilizar datos desalineados o calcular una medida inapropiada. La depuración debe examinar tanto el funcionamiento técnico como la correspondencia entre el código y el procedimiento científico previsto.

La visualización asistida permite convertir resultados en gráficos, mapas y diagramas que facilitan reconocer tendencias, diferencias y valores atípicos. El sistema puede sugerir representaciones según el tipo de datos, pero la elección final debe considerar claridad, proporcionalidad y fidelidad. Los gráficos no son elementos decorativos, sino formas de argumentación científica capaces de destacar u ocultar relaciones. El investigador debe revisar escalas, etiquetas, intervalos y criterios de agrupación para evitar interpretaciones engañosas, especialmente cuando la figura resume conjuntos complejos o compara poblaciones con tamaños diferentes.

La comprobación de supuestos estadísticos continúa siendo indispensable aunque el análisis haya sido generado automáticamente. Deben evaluarse independencia, distribución, homogeneidad, linealidad, multicolinealidad y otros requisitos según el procedimiento utilizado. Barberis et al. (2024) subrayan la importancia de aplicar prácticas estandarizadas para fortalecer la robustez y reproducibilidad del aprendizaje automático en ciencias biomédicas. Cuando los supuestos no se cumplen, el investigador debe justificar transformaciones, métodos alternativos o interpretaciones restringidas. Omitir estas comprobaciones puede producir conclusiones precisas en apariencia, pero metodológicamente débiles.

La programación asistida debe organizarse como un ciclo de solicitud, ejecución, diagnóstico y validación, no como una aceptación inmediata del código producido. La inteligencia artificial puede ampliar la autonomía técnica y acelerar análisis, pero la responsabilidad metodológica permanece en el equipo investigador. Cada salida debe compararse con resultados esperados, procedimientos alternativos y conocimiento disciplinar. La revisión experta permite determinar si el análisis responde realmente a la pregunta, si conserva la integridad de los datos y si sus conclusiones pueden comunicarse con el grado de incertidumbre apropiado.

Figura 9

Ciclo de análisis estadístico y programación asistida por inteligencia artificial



Nota. La figura representa un proceso iterativo de programación y análisis estadístico asistido por inteligencia artificial. El código generado debe ejecutarse, diagnosticarse y validarse antes de incorporarse a las conclusiones científicas.

El análisis estadístico y la programación asistidos por inteligencia artificial deben desarrollarse como un proceso iterativo de generación, ejecución, diagnóstico, corrección y validación. El código producido constituye una propuesta inicial que puede contener errores técnicos, supuestos inadecuados o decisiones metodológicas poco pertinentes. Por ello, sus resultados deben compararse con métodos alternativos, datos independientes y conocimiento disciplinar. La documentación de versiones, parámetros y modificaciones permite conservar la trazabilidad y garantiza que las conclusiones científicas se apoyen en procedimientos revisados y reproducibles.

3.4 Laboratorios autónomos y experimentación robótica

Los laboratorios autónomos integran inteligencia artificial, robótica, instrumentación científica y sistemas de decisión para ejecutar experimentos con una intervención humana reducida. Su desarrollo responde a la necesidad de explorar espacios experimentales demasiado amplios, costosos o lentos para ser examinados manualmente. Abolhasani y Kumacheva (2023) explican que estos entornos articulan automatización física, aprendizaje automático y planificación experimental. Su finalidad no consiste únicamente en acelerar tareas, sino en seleccionar procedimientos, interpretar resultados parciales y orientar nuevas acciones dentro de límites previamente definidos por los investigadores.

Es necesario diferenciar automatización y autonomía. Un laboratorio automatizado ejecuta secuencias programadas, repite operaciones y controla instrumentos según instrucciones fijas. En cambio, un sistema autónomo modifica sus decisiones a partir de los resultados obtenidos y selecciona nuevas condiciones experimentales. Xie et al. (2023) señalan que la convergencia entre inteligencia artificial y automatización experimental permite avanzar hacia infraestructuras capaces de percibir, decidir y actuar. La autonomía, por tanto, implica adaptación, aunque sus objetivos, restricciones y criterios de seguridad continúan siendo establecidos por personas.

La robótica proporciona la capacidad física necesaria para manipular muestras, trasladar materiales, dosificar sustancias y controlar condiciones experimentales. Los robots pueden ejecutar tareas repetitivas con precisión constante y reducir variaciones asociadas con procedimientos manuales. Sin embargo, cada operación requiere calibración, protocolos definidos y mecanismos para detectar fallas. Tom et al. (2024) destacan que los laboratorios autónomos dependen de la integración entre equipos físicos y sistemas

digitales. Una decisión computacional carece de utilidad si la infraestructura robótica no puede ejecutarla de manera segura y reproducible.

Los sensores y los instrumentos científicos constituyen la base de percepción del laboratorio. Mediante ellos se registran temperatura, presión, composición, imágenes, señales espectroscópicas y otras propiedades relevantes. Estos datos permiten determinar si el experimento se desarrolló según lo previsto y proporcionan evidencia para actualizar el modelo. La calidad de la autonomía depende directamente de la precisión instrumental, porque mediciones erróneas pueden orientar decisiones posteriores hacia regiones poco prometedoras. Por ello, la calibración, el mantenimiento y el control de calidad permanecen como condiciones esenciales del sistema experimental.

La planificación computacional organiza objetivos, recursos, secuencias y restricciones antes de ejecutar cada procedimiento. El sistema puede consultar conocimiento previo, comparar rutas y seleccionar aquella que maximiza la información esperada o la probabilidad de éxito. Boiko et al. (2023) demostraron que un modelo de lenguaje conectado con herramientas especializadas puede planificar procedimientos químicos y coordinar diferentes componentes. Esta capacidad muestra que la planificación deja de ser una lista estática de instrucciones y se convierte en una actividad que puede modificarse según el contexto experimental disponible.

El aprendizaje activo permite seleccionar el siguiente experimento considerando aquello que el modelo todavía desconoce. En lugar de probar combinaciones al azar, el sistema prioriza condiciones capaces de reducir incertidumbre o mejorar una propiedad objetivo. Abolhasani y Kumacheva (2023) identifican esta estrategia como uno de los componentes centrales de los laboratorios autónomos. Su ventaja consiste en obtener mayor información con menos experimentos. No obstante, la utilidad del aprendizaje activo depende de cómo se define la función objetivo y de qué variables quedan representadas en el modelo.

La selección experimental también debe equilibrar explotación y exploración. La explotación concentra recursos en regiones donde el modelo espera buenos resultados, mientras la exploración evalúa alternativas menos conocidas que podrían revelar comportamientos inesperados. Un sistema excesivamente orientado a la explotación puede repetir soluciones previsibles y pasar por alto descubrimientos novedosos. Por el contrario, una exploración demasiado amplia puede consumir tiempo y materiales sin

producir avances. El equilibrio entre ambas estrategias debe ajustarse según el propósito científico, la disponibilidad de recursos y el nivel de incertidumbre aceptable.

La ejecución robótica transforma la propuesta computacional en una intervención física. Para ello, las instrucciones deben convertirse en movimientos, cantidades, tiempos y condiciones operativas comprensibles para los equipos. Szymanski et al. (2023) integraron planificación computacional, aprendizaje activo e instrumentación automatizada en A-Lab para sintetizar materiales inorgánicos. Este caso demuestra que la autonomía requiere compatibilidad entre modelos, protocolos y dispositivos. Cuando la traducción entre niveles falla, pueden producirse errores experimentales, aunque la propuesta generada por el sistema sea conceptualmente adecuada.

La medición posterior determina si el resultado obtenido corresponde con el objetivo planteado. Los instrumentos analizan productos, propiedades y posibles desviaciones, generando datos que regresan al sistema. Esta fase permite distinguir entre una ejecución técnicamente correcta y un experimento científicamente exitoso. Tom et al. (2024) señalan que la caracterización automatizada constituye un componente indispensable para vincular experimentación y toma de decisiones. Sin mediciones confiables, el laboratorio solo automatiza procedimientos, pero no puede aprender de sus resultados ni modificar racionalmente las acciones siguientes.

La retroalimentación conecta medición, interpretación y nueva decisión. El sistema actualiza sus estimaciones, reevalúa alternativas y propone ajustes en las condiciones experimentales. Este mecanismo incrementa la capacidad de adaptación y permite corregir trayectorias que no producen resultados útiles. Sin embargo, la retroalimentación también puede amplificar errores cuando las mediciones son defectuosas o el modelo interpreta incorrectamente la evidencia. Por esta razón, deben existir umbrales de revisión, alertas y criterios que obliguen a detener la operación cuando los resultados se apartan de rangos seguros o científicamente plausibles.

Los principales casos de aplicación se concentran en química y ciencia de materiales, campos donde existe una gran cantidad de combinaciones posibles. Boiko et al. (2023) mostraron la coordinación automatizada de procedimientos químicos, mientras Szymanski et al. (2023) aplicaron autonomía experimental a la síntesis de materiales. Xie et al. (2023) destacan que estos sistemas pueden acelerar descubrimientos al unir diseño computacional y experimentación física. No obstante, la transferencia hacia otras

disciplinas exige adaptar instrumentos, representaciones de datos, protocolos y criterios de validación a cada objeto de estudio.

Un laboratorio autónomo debe entenderse como una arquitectura integrada y no como un único robot o modelo. Su funcionamiento depende de planificación, sensores, instrumentación, ejecución, medición y actualización coordinadas. La autonomía efectiva requiere que cada componente sea trazable, seguro y compatible con los demás. La supervisión humana conserva la definición de objetivos, la autorización de operaciones críticas y la evaluación final de los resultados. La automatización física amplía la capacidad experimental, pero la validez científica continúa dependiendo de procedimientos reproducibles y decisiones justificadas.

Tabla 10

Casos representativos de laboratorios autónomos y experimentación robótica

Sistema o estudio	Disciplina	Nivel de autonomía	Componentes principales	Resultado científico
Coscientist	Química	Autonomía parcial con herramientas conectadas	Modelo de lenguaje, búsqueda, código e instrumentación	Planificación y ejecución de procedimientos químicos
A-Lab	Ciencia de materiales	Autonomía experimental elevada	Modelos computacionales, aprendizaje activo, robótica y caracterización	Síntesis y comprobación de materiales inorgánicos
Laboratorios autodirigidos descritos por Abolhasani y Kumacheva	Química y materiales	Variable según infraestructura	Automatización, sensores, aprendizaje automático y planificación	Optimización acelerada de condiciones experimentales
Arquitecturas revisadas por Xie et al.	Ciencia de materiales	Desde automatización avanzada hasta autonomía adaptativa	Diseño experimental, robots, análisis y toma de decisiones	Reducción del espacio experimental y selección inteligente de pruebas
Sistemas revisados por Tom et al.	Química y materiales	Diversos niveles de madurez	Hardware, software, gestión de datos y aprendizaje activo	Integración de experimentación física y decisión computacional

Nota. La tabla compara sistemas que combinan automatización física y decisión computacional en diferentes grados. El nivel de autonomía depende de la capacidad para modificar la selección experimental a partir de las mediciones obtenidas.

Los casos presentados muestran que los laboratorios autónomos combinan distintos grados de automatización física, análisis computacional y toma de decisiones adaptativa. Su nivel de autonomía aumenta cuando el sistema puede seleccionar nuevos experimentos a partir de los resultados obtenidos, en lugar de limitarse a ejecutar secuencias predefinidas. Aunque estas arquitecturas aceleran la planificación, la síntesis y la optimización experimental, su validez depende de la integración técnica, la calidad de las mediciones, la trazabilidad de las decisiones y la supervisión humana sobre las operaciones críticas.

Figura 10
Arquitectura general de un laboratorio autónomo



Nota. La figura representa los componentes técnicos de un laboratorio autónomo y su interacción. La autonomía surge cuando los resultados medidos modifican la selección de experimentos posteriores, mientras que la supervisión humana conserva el control de los objetivos y operaciones críticas.

Un laboratorio autónomo funciona mediante la integración coordinada de planificación, aprendizaje activo, instrumentación, medición y análisis computacional. La autonomía se manifiesta cuando los resultados obtenidos modifican la selección de los experimentos posteriores y permiten ajustar progresivamente las condiciones de trabajo. Sin embargo, este proceso requiere límites operativos, controles de seguridad y criterios de validación claramente definidos. La supervisión humana sigue siendo indispensable para establecer objetivos, autorizar decisiones críticas y garantizar que la eficiencia tecnológica no comprometa el rigor científico ni la seguridad experimental.

3.5 Automatización de análisis cualitativo y multimodal

La automatización del análisis cualitativo y multimodal permite procesar grandes volúmenes de textos, imágenes, audios, videos y datos estructurados dentro de un mismo flujo investigativo. Su aporte principal consiste en acelerar tareas de organización, codificación preliminar, comparación y reconocimiento de relaciones entre modalidades. Goyanes y Lopezosa (2024) muestran que la inteligencia artificial puede apoyar entrevistas y análisis de contenido, aunque sus resultados requieren supervisión. En este contexto, automatizar no significa sustituir la interpretación, sino ampliar la capacidad para examinar materiales complejos y heterogéneos.

En los datos textuales, los sistemas pueden segmentar documentos, identificar expresiones recurrentes y proponer códigos iniciales. Estas funciones facilitan la revisión de entrevistas, respuestas abiertas, informes y publicaciones digitales. Perkins y Roe (2024) explican que los modelos generativos pueden apoyar el análisis temático inductivo cuando reciben instrucciones claras y trabajan dentro de un procedimiento definido. Sin embargo, las categorías sugeridas deben contrastarse con los fragmentos originales, porque la semejanza lingüística no siempre representa equivalencia conceptual ni conserva la intención expresada por los participantes.

La automatización también puede comparar discursos entre personas, grupos o periodos. Los modelos permiten localizar coincidencias, contradicciones y cambios en el uso de determinados conceptos, generando matrices que organizan el corpus. Nguyen (2025) señala que la inteligencia artificial puede desempeñarse como asistente durante el análisis temático, especialmente en tareas de exploración y clasificación. El investigador debe comprobar si las diferencias identificadas responden a posiciones sociales, circunstancias

históricas o variaciones contextuales, evitando que el análisis se limite a contabilizar palabras o proximidades semánticas.

En el análisis de imágenes, la inteligencia artificial puede reconocer objetos, formas, expresiones, escenas y características visuales. Estas capacidades resultan útiles para examinar fotografías, materiales educativos, documentos escaneados o registros clínicos. La automatización permite etiquetar grandes colecciones y localizar patrones que luego pueden analizarse cualitativamente. Sin embargo, el contenido visual no posee un significado universal. Una imagen adquiere sentido según su origen, propósito, audiencia y condiciones de producción, elementos que no siempre se encuentran representados en los datos utilizados por el modelo.

El procesamiento de audio puede incluir transcripción, identificación de hablantes, reconocimiento de emociones aparentes y segmentación de intervenciones. En video, además, se integran movimiento, secuencia temporal, imagen y sonido. Estas modalidades generan información más rica, pero también aumentan las posibilidades de error. Los gestos, silencios y entonaciones pueden variar culturalmente y no deben traducirse automáticamente en categorías psicológicas o sociales. Girdwood y Manikonda (2026) advierten que las decisiones prácticas y éticas deben acompañar todo el proceso cualitativo asistido por inteligencia artificial.

Los datos clínicos ofrecen un ejemplo relevante de análisis multimodal porque pueden combinar imágenes, resultados de laboratorio, historias médicas, señales biológicas y texto especializado. Truhn et al. (2024) sostienen que los modelos fundacionales multimodales pueden integrar estas fuentes para apoyar investigaciones en oncología de precisión. La combinación permite construir representaciones más completas que aquellas derivadas de una sola modalidad. No obstante, cada fuente posee niveles diferentes de calidad, temporalidad y precisión, por lo que su integración debe considerar compatibilidad, procedencia y posibles discrepancias entre registros.

La automatización multimodal requiere transformar datos distintos en representaciones que puedan analizarse conjuntamente. El sistema puede relacionar una imagen con su descripción, un fragmento de audio con una transcripción o una medición clínica con una nota textual. Bennis y Mouwafaq (2025) muestran que la inteligencia artificial puede apoyar el análisis temático en conjuntos cualitativos amplios. Al integrar modalidades, el

investigador debe comprobar si las relaciones propuestas reflejan conexiones reales o si surgen de patrones estadísticos aprendidos en contextos distintos del estudio desarrollado.

Uno de los riesgos principales es la pérdida de contexto durante la conversión y combinación de datos. Una transcripción elimina parte de la expresión oral; una etiqueta visual reduce una escena a categorías; y un resumen puede omitir contradicciones relevantes. La automatización tiende a favorecer información repetida, clara y fácilmente clasificable, mientras puede ignorar ambigüedades, silencios o casos excepcionales. Por ello, los resultados automatizados deben vincularse con sus materiales originales, permitiendo que el investigador regrese al audio, imagen, video o documento para revisar el significado atribuido.

Los sesgos culturales también afectan la interpretación automatizada. Los modelos pueden estar entrenados con idiomas, imágenes y comportamientos procedentes de poblaciones específicas, generando errores cuando analizan contextos diferentes. Una expresión puede clasificarse de forma incorrecta, un gesto recibir un significado ajeno a la cultura estudiada o una variedad lingüística ser considerada inconsistente. Girdwood y Manikonda (2026) resaltan la necesidad de evaluar cómo la IA interviene desde el diseño hasta la publicación. La revisión contextual debe incluir diversidad lingüística, social y geográfica.

La triangulación permite contrastar los resultados producidos desde distintas modalidades y procedimientos. Una categoría obtenida del texto puede compararse con imágenes, observaciones, registros cuantitativos o interpretaciones de otros investigadores. También resulta útil contrastar la codificación automatizada con análisis manuales y revisar casos donde ambos procedimientos divergen. Goyanes y Lopezosa (2024) señalan que la colaboración humano-máquina debe sostenerse en pensamiento crítico. La coincidencia fortalece una interpretación cuando existe evidencia suficiente, mientras el desacuerdo puede revelar errores o dimensiones analíticas nuevas.

En consecuencia, la automatización cualitativa y multimodal debe organizarse como un proceso de asistencia verificable. La inteligencia artificial facilita la codificación, comparación e integración de materiales, pero el significado permanece condicionado por preguntas, contextos y marcos interpretativos. La supervisión humana debe conservar la conexión con los datos originales, reconocer sesgos y justificar cómo se integraron las modalidades. El valor científico no reside en procesar más información, sino en construir

interpretaciones coherentes, trazables y sensibles a la diversidad de formas en que los fenómenos fueron registrados.

Tabla 11

Automatización del análisis cualitativo y multimodal

Modalidad	Tipo de dato	Proceso automatizable	Resultado posible	Validación humana necesaria
Texto	Entrevistas, documentos, respuestas abiertas y publicaciones	Segmentación, codificación inicial, búsqueda de recurrencias y agrupamiento	Códigos, categorías preliminares y relaciones temáticas	Revisar significado, contexto, coherencia y fragmentos originales
Imagen	Fotografías, gráficos, documentos y registros clínicos	Reconocimiento de objetos, escenas, formas y características visuales	Etiquetas, agrupaciones y patrones visuales	Interpretar origen, propósito, contexto cultural y límites de la clasificación
Audio	Entrevistas, conversaciones y registros sonoros	Transcripción, separación de hablantes y segmentación	Texto organizado, turnos de habla y marcas temporales	Corregir errores y conservar entonación, silencios y expresiones relevantes
Video	Movimiento, imagen, sonido e interacción	Identificación de secuencias, acciones y eventos	Patrones de comportamiento y escenas clasificadas	Analizar temporalidad, relaciones sociales y contexto situacional
Datos clínicos	Imágenes, señales, historias médicas y resultados de laboratorio	Integración de modalidades y reconocimiento de relaciones	Perfiles, clasificaciones y representaciones integradas	Comprobar procedencia, compatibilidad, calidad y pertinencia clínica
Datos estructurados	Tablas, variables, registros y metadatos	Normalización, vinculación y comparación con información cualitativa	Matrices integradas y asociaciones entre fuentes	Verificar equivalencias, unidades, temporalidad y correspondencia entre casos
Análisis multimodal	Combinación de texto, imagen,	Construcción de representaciones	Interpretaciones integradas y	Contrastar cada modalidad por

Modalidad	Tipo de dato	Proceso automatizable	Resultado posible	Validación humana necesaria
	audio, video y datos numéricos	conjuntas y búsqueda de correspondencias	relaciones entre modalidades	separado y justificar su integración
Triangulación	Resultados automatizados, análisis manuales y fuentes complementarias	Comparación de coincidencias y discrepancias	Confirmación, reformulación o ampliación de categorías	Valorar evidencia contradictoria y evitar asumir consenso automático

Nota. La tabla presenta las principales modalidades susceptibles de análisis automatizado y los controles requeridos para conservar contexto, significado y trazabilidad. La integración multimodal debe complementar, no reemplazar, la interpretación humana.

La automatización del análisis cualitativo y multimodal permite organizar, comparar e integrar textos, imágenes, audios, videos y datos estructurados con mayor rapidez. Sin embargo, los resultados obtenidos conservan un carácter preliminar, porque la clasificación automática puede perder contexto, simplificar significados o establecer relaciones poco justificadas entre modalidades. La validación humana debe revisar las fuentes originales, interpretar las condiciones de producción y valorar las discrepancias. De este modo, la inteligencia artificial complementa el análisis sin sustituir la comprensión contextual ni la construcción reflexiva de las conclusiones.

3.6 Redacción, visualización y comunicación asistidas

La inteligencia artificial puede apoyar la comunicación científica desde la organización inicial del manuscrito hasta la adaptación de los resultados para distintos públicos. Su intervención abarca estructuración, revisión lingüística, síntesis, traducción y elaboración preliminar de recursos visuales. Lopezosa (2023) sostiene que estas tecnologías pueden mejorar la comunicación académica cuando se utilizan de manera útil y responsable. Sin embargo, el apoyo automatizado no sustituye la comprensión del estudio, porque la coherencia entre objetivos, métodos, resultados y conclusiones depende de decisiones intelectuales asumidas por los investigadores.

Durante la estructuración del manuscrito, los modelos generativos pueden proponer esquemas, ordenar apartados y detectar contenidos ubicados fuera de la sección correspondiente. Este apoyo resulta pertinente cuando el equipo necesita convertir materiales dispersos en una secuencia argumentativa comprensible. Román (2023)

destaca que la inteligencia artificial puede facilitar tareas de investigación académica asociadas con organización y producción textual. No obstante, la estructura sugerida debe contrastarse con el género documental, las normas editoriales y la lógica metodológica, evitando utilizar plantillas genéricas que fragmenten o repitan innecesariamente la información.

La revisión lingüística comprende corrección gramatical, puntuación, claridad sintáctica y reducción de redundancias. Los sistemas pueden señalar frases ambiguas, transiciones débiles y variaciones terminológicas que afectan la lectura. Esta asistencia debe conservar el significado científico y la voz de los autores, porque una reformulación aparentemente elegante puede modificar el alcance de una afirmación. La corrección automatizada tampoco debe eliminar conceptos técnicos necesarios ni reemplazar expresiones disciplinares por términos imprecisos. El texto final necesita una lectura experta capaz de comprobar precisión conceptual, cohesión y correspondencia con los resultados.

La traducción asistida facilita la circulación internacional de artículos, informes y materiales divulgativos. Los modelos pueden producir versiones preliminares y adaptar expresiones según el registro académico, pero deben revisarse cuidadosamente nombres de variables, términos especializados y relaciones lógicas. Gimeno y Trigo (2024) analizan el papel creciente de la inteligencia artificial en la publicación científica y reconocen posibilidades de apoyo editorial. La calidad de una traducción no depende solo de equivalencias lingüísticas, sino de conservar el sentido metodológico y la terminología aceptada dentro de cada comunidad científica.

La síntesis automatizada permite elaborar resúmenes, puntos destacados, títulos y palabras clave a partir del contenido del manuscrito. Esta función puede ayudar a condensar resultados y adaptar la extensión a los requisitos de una revista. Sin embargo, los modelos tienden a privilegiar afirmaciones generales o llamativas, por lo que podrían omitir limitaciones y matices esenciales. Montomoli (2023) advierte que la inteligencia artificial está modificando la publicación científica, especialmente en tareas editoriales. Cada resumen debe comprobarse contra el texto completo para evitar exageraciones, omisiones o conclusiones no sustentadas.

La generación de tablas y gráficos puede automatizarse mediante instrucciones que indiquen variables, relaciones y formatos. Los sistemas también pueden producir código para visualizar distribuciones, tendencias o comparaciones. Shao et al. (2025) muestran

que los colaboradores basados en inteligencia artificial pueden agilizar flujos analíticos y facilitar la generación de productos científicos. Pese a ello, el investigador debe verificar datos, escalas, etiquetas y criterios de agrupación. Una visualización técnicamente correcta puede inducir interpretaciones equivocadas si utiliza proporciones inadecuadas o destaca relaciones irrelevantes.

Los diagramas y figuras conceptuales permiten representar procesos, arquitecturas y relaciones difíciles de comunicar mediante texto continuo. La inteligencia artificial puede proponer composiciones, jerarquías y conexiones, pero el recurso debe cumplir una función explicativa real. Las imágenes decorativas de robots, cerebros digitales o circuitos no aportan evidencia ni comprensión metodológica. Antes de incorporar una figura, conviene preguntar qué relación permite observar, qué información sintetiza y cómo complementa el texto. También deben revisarse símbolos, secuencias y conceptos para evitar representaciones visualmente atractivas, pero científicamente incorrectas.

La comunicación asistida puede adaptar un mismo resultado para artículos, presentaciones, informes técnicos, materiales docentes o divulgación pública. Lu et al. (2026) presentan sistemas capaces de participar en múltiples fases, incluida la elaboración de manuscritos y la evaluación automatizada. Esta flexibilidad permite modificar vocabulario, profundidad y extensión según la audiencia. Sin embargo, adaptar no significa simplificar hasta perder exactitud. La comunicación dirigida a públicos no especializados debe explicar incertidumbres y límites sin introducir afirmaciones absolutas que el estudio original no permite sostener.

La verificación final debe abarcar cifras, citas, referencias, tablas, gráficos y correspondencia entre texto y evidencia. Ninguna fuente debe incorporarse únicamente porque fue sugerida por un sistema; es necesario confirmar autores, título, año, DOI y contenido original. Asimismo, el uso de inteligencia artificial debe declararse cuando haya intervenido sustancialmente en la elaboración o transformación del material. La responsabilidad del autor comprende revisar cada resultado, justificar las decisiones comunicativas y garantizar que el manuscrito representa fielmente el proceso científico desarrollado, sin delegar su exactitud en la herramienta utilizada.

Tabla 12

Usos de la inteligencia artificial en la redacción y comunicación científica

Tarea comunicativa	Ayuda permitida de la inteligencia artificial	Verificación necesaria	Responsabilidad del autor
Estructuración del manuscrito	Proponer esquemas, secciones y secuencias argumentativas	Comprobar correspondencia con objetivos, métodos y género documental	Decidir la organización definitiva del texto
Revisión lingüística	Corregir gramática, puntuación, claridad y redundancias	Confirmar que no cambie el significado científico	Garantizar precisión conceptual y coherencia
Traducción	Generar versiones preliminares en otros idiomas	Revisar terminología, variables, relaciones lógicas y registro académico	Aprobar la versión final y asumir su exactitud
Elaboración de resúmenes	Condensar objetivos, métodos, resultados y conclusiones	Comparar con el manuscrito completo y recuperar limitaciones relevantes	Evitar exageraciones u omisiones sustantivas
Títulos y palabras clave	Sugerir alternativas y términos de indexación	Verificar pertinencia temática y correspondencia con el contenido	Seleccionar opciones claras, precisas y representativas
Tablas	Organizar datos y comparaciones en formatos estructurados	Comprobar cifras, categorías, totales, títulos y notas	Asegurar fidelidad respecto de los resultados
Gráficos	Generar código y proponer visualizaciones	Revisar escalas, etiquetas, proporciones y correspondencia con los datos	Elegir la representación adecuada y evitar distorsiones
Diagramas y figuras	Representar procesos, relaciones y arquitecturas	Validar secuencia, conceptos, símbolos y conexiones	Garantizar que la figura tenga valor explicativo
Divulgación	Adaptar vocabulario, extensión y nivel técnico	Confirmar que la simplificación conserve exactitud y limitaciones	Ajustar el mensaje al público sin alterar la evidencia
Revisión de referencias	Detectar inconsistencias y completar datos preliminares	Verificar cada referencia en la fuente original	Responder por la existencia, exactitud y pertinencia de las citas

Tarea comunicativa	Ayuda permitida de la inteligencia artificial	Verificación necesaria	Responsabilidad del autor
Declaración del uso	Ayudar a redactar una nota metodológica sobre la intervención de IA	Precisar herramienta, función y alcance real del apoyo	Informar de manera transparente cómo fue utilizada

Nota. La tabla distingue tareas de redacción, visualización y comunicación en las que la inteligencia artificial puede brindar apoyo. Toda salida requiere revisión humana para garantizar fidelidad, claridad, trazabilidad y correspondencia con la evidencia científica.

La inteligencia artificial puede apoyar la redacción y comunicación científica mediante la organización del manuscrito, la revisión lingüística, la traducción, la elaboración de resúmenes y la creación de recursos visuales. No obstante, toda intervención debe someterse a verificación para evitar errores, omisiones o alteraciones del significado científico. La responsabilidad final corresponde al autor, quien debe garantizar la exactitud de las referencias, la fidelidad de tablas y gráficos, la pertinencia del lenguaje y la declaración transparente del uso de estas herramientas.

3.7 Agentes autónomos y automatización integral del ciclo científico

Un agente autónomo es un sistema de inteligencia artificial capaz de recibir un objetivo, planificar acciones, utilizar herramientas y ajustar su comportamiento según los resultados obtenidos. A diferencia de un modelo generativo que responde a una instrucción aislada, el agente mantiene una secuencia de trabajo y decide qué operación realizar después. Boiko et al. (2023) mostraron que un modelo conectado con búsqueda, programación e instrumentación puede coordinar tareas científicas complejas. Sin embargo, esta autonomía permanece delimitada por objetivos, permisos, recursos y controles establecidos previamente por investigadores humanos.

La memoria constituye un componente central porque permite conservar información sobre decisiones, datos consultados, resultados parciales y errores encontrados durante el proceso. Puede adoptar la forma de historial de interacciones, registros estructurados, documentos recuperados o bases de conocimiento actualizadas. Sin memoria, el agente repite tareas y pierde continuidad entre etapas. No obstante, almacenar información no garantiza que esta sea correcta ni relevante. Los registros deben conservar procedencia, fecha y condiciones de uso para evitar que una conclusión provisional se transforme en supuesto permanente durante las operaciones posteriores.

La planificación permite dividir un objetivo amplio en actividades más pequeñas y ordenarlas según dependencias, recursos y prioridades. Un agente puede decidir comenzar con la exploración bibliográfica, continuar con la formulación de hipótesis y luego generar código para analizar datos. Lu et al. (2026) describen un sistema orientado a automatizar etapas sucesivas de investigación en inteligencia artificial. Este tipo de planificación amplía la capacidad operativa, aunque puede desviarse cuando los objetivos son ambiguos o los criterios de éxito no están suficientemente definidos por el equipo investigador.

El uso de herramientas convierte al agente en una interfaz capaz de actuar sobre entornos externos. Puede consultar repositorios, ejecutar programas, analizar archivos, llamar modelos especializados o enviar instrucciones a instrumentos científicos. Boiko et al. (2023) evidencian esta capacidad en procedimientos químicos coordinados mediante distintos componentes. Cada herramienta introduce posibilidades y riesgos específicos. Una búsqueda incompleta, un código defectuoso o una medición incorrecta pueden afectar toda la secuencia. Por ello, las conexiones deben limitar permisos, registrar operaciones y verificar la integridad de los resultados transferidos entre componentes.

Las arquitecturas multiagente distribuyen funciones entre sistemas especializados. Un agente puede explorar literatura, otro generar hipótesis, un tercero analizar datos y otro revisar críticamente las conclusiones. Ghareeb et al. (2026) proponen una organización de este tipo para automatizar actividades vinculadas con el descubrimiento científico. La especialización permite dividir problemas complejos y comparar perspectivas. Sin embargo, el número de agentes no garantiza mayor rigor. Todos pueden depender de fuentes semejantes, reproducir supuestos compartidos o aceptar de forma colectiva una afirmación incorrecta si no existen mecanismos independientes de comprobación.

La coordinación determina cómo se comparten tareas, resultados y desacuerdos entre agentes. Puede existir una estructura jerárquica, donde un coordinador asigna funciones, o una organización colaborativa basada en intercambio y evaluación mutua. Gottweis et al. (2026) muestran una arquitectura en la que distintos agentes generan, critican, clasifican y refinan hipótesis científicas. Este debate puede mejorar la calidad de las propuestas, aunque necesita reglas para evitar ciclos interminables, consensos superficiales o predominio de argumentos persuasivos sobre evidencias. La decisión final debe permanecer vinculada con criterios científicos explícitos.

En la fase de ideación, los agentes pueden combinar literatura, datos y conocimiento previo para formular preguntas o hipótesis. Cornelio et al. (2023) demuestran que los sistemas computacionales pueden descubrir relaciones cuando integran evidencia y principios teóricos. Una arquitectura autónoma puede ampliar esta capacidad mediante búsquedas iterativas y comparación de alternativas. No obstante, generar muchas ideas no equivale a producir contribuciones significativas. Las propuestas deben priorizarse según novedad, plausibilidad, relevancia, viabilidad y comprobabilidad, evitando que la cantidad de alternativas sustituya la valoración disciplinar de su importancia.

La programación asistida permite que agentes especializados generen código, ejecuten análisis y modifiquen procedimientos según diagnósticos intermedios. Lu et al. (2026) integran generación de ideas, desarrollo de código, experimentación computacional y elaboración de manuscritos dentro de un flujo automatizado. Esta capacidad acelera la comparación entre configuraciones, pero también favorece la propagación de errores. Un código que funciona puede utilizar variables incorrectas, aplicar una métrica inadecuada o interpretar de forma errónea los resultados. Cada experimento necesita controles técnicos, metodológicos y documentales antes de alimentar nuevas decisiones.

En entornos experimentales, los agentes pueden seleccionar condiciones, coordinar instrumentos y analizar mediciones para decidir qué procedimiento ejecutar después. Boiko et al. (2023) muestran esta conexión entre modelos de lenguaje y experimentación química, mientras Cory et al. (2024) evidencian cómo el conocimiento teórico puede orientar búsquedas computacionales. La automatización integral exige que planificación digital y ejecución física permanezcan sincronizadas. Los límites de seguridad, materiales disponibles, calibración instrumental y criterios de interrupción deben establecerse antes de permitir que el sistema modifique autónomamente una secuencia experimental.

El análisis y la redacción pueden integrarse dentro del mismo flujo. Los agentes organizan resultados, producen visualizaciones, comparan hipótesis y redactan versiones preliminares de informes o artículos. Esta continuidad reduce la separación entre ejecución y comunicación, pero introduce el riesgo de que una interpretación incorrecta se replique en tablas, argumentos y conclusiones. El sistema debe distinguir datos originales, resultados calculados e inferencias elaboradas. La revisión humana necesita comprobar cada transición, porque una narrativa coherente puede ocultar fallas metodológicas acumuladas durante las etapas anteriores del proceso científico.

La evaluación automática incorpora agentes encargados de revisar métodos, resultados, código y consistencia argumentativa. Tang et al. (2025) advierten que el avance hacia científicos de inteligencia artificial plantea riesgos que requieren priorizar salvaguardas sobre autonomía. Un evaluador automático puede detectar contradicciones o debilidades, pero no debe considerarse completamente independiente cuando comparte arquitectura, datos o criterios con el agente evaluado. La confiabilidad aumenta mediante verificaciones externas, pruebas adversariales, revisión experta y separación entre generación y evaluación, especialmente antes de ejecutar acciones irreversibles o comunicar hallazgos.

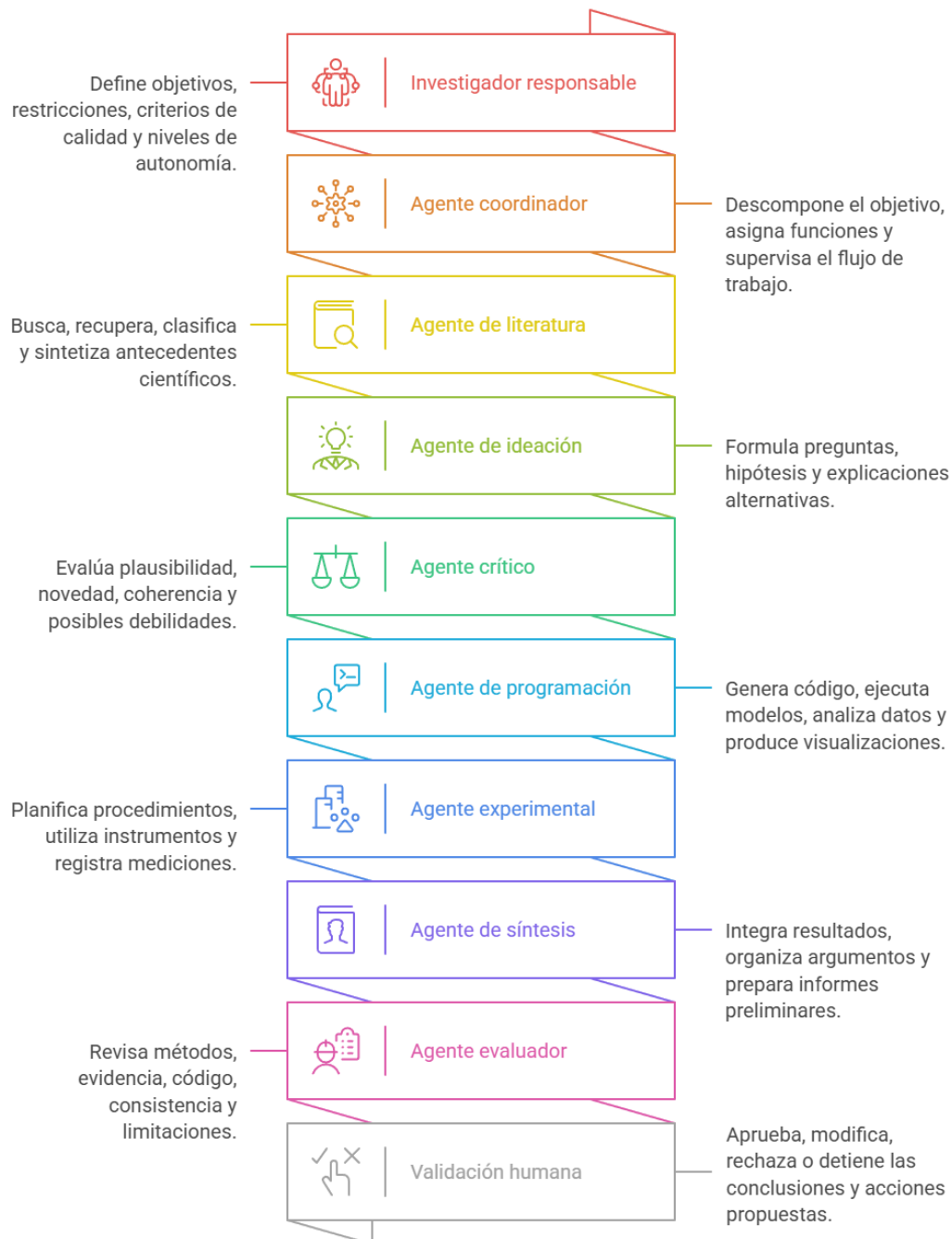
Por tanto, la automatización integral del ciclo científico debe entenderse como autonomía gradual y supervisada, no como sustitución consolidada del investigador. Los agentes actuales pueden conectar ideación, búsqueda, programación, experimentación, análisis, redacción y evaluación, pero su desempeño depende de la calidad de cada componente y de la coordinación entre ellos. La responsabilidad humana sigue siendo necesaria para definir objetivos, aprobar procedimientos críticos, validar resultados y detener procesos inseguros. El progreso científico no se mide por cuántas tareas se delegan, sino por cuánto rigor, trazabilidad y comprensión conserva el sistema completo.

La automatización integral del ciclo científico requiere una organización capaz de distribuir funciones entre agentes especializados. En lugar de concentrar todas las tareas en un único sistema, la arquitectura multiagente asigna responsabilidades diferenciadas para buscar literatura, formular hipótesis, generar código, ejecutar análisis, planificar experimentos y revisar resultados. Esta división favorece la especialización y permite que cada agente opere con objetivos, herramientas y criterios específicos, mientras un coordinador administra las dependencias y mantiene la continuidad del proceso.

Sin embargo, la colaboración entre agentes no garantiza por sí sola la calidad científica. Los errores pueden propagarse desde una búsqueda incompleta hasta la interpretación, redacción o evaluación final si no existen controles adecuados. Por ello, la arquitectura debe incorporar memoria compartida, trazabilidad, permisos graduados, registros de versiones y verificación externa. La supervisión humana permanece como elemento transversal, pues define los objetivos, interviene ante discrepancias y conserva la facultad de modificar, detener o rechazar las acciones propuestas.

Figura 11

Arquitectura multiagente para la automatización integral del ciclo científico



Nota. La figura representa una arquitectura multiagente en la que funciones especializadas participan en diferentes etapas del ciclo científico. La coordinación automatizada no elimina la supervisión humana, que conserva el control sobre objetivos, procedimientos críticos, validación y comunicación de resultados.

La arquitectura multiagente permite coordinar capacidades especializadas dentro de un proceso científico integrado, sin convertir la automatización en autonomía irrestricta. Su utilidad depende de que las funciones estén claramente delimitadas, las decisiones puedan reconstruirse y los resultados sean sometidos a evaluación independiente. La intervención humana continúa siendo indispensable para validación de la información.

Capítulo

4

Calidad y transparencia en investigaciones con inteligencia artificial

Jefferson Paul Sánchez Vivanco



4.1 Validez y confiabilidad de resultados asistidos por inteligencia artificial

La incorporación de inteligencia artificial en la investigación exige evaluar si los resultados obtenidos representan adecuadamente el fenómeno estudiado y si pueden mantenerse bajo condiciones comparables. La validez se refiere a la correspondencia entre las conclusiones, los datos y la pregunta científica, mientras que la confiabilidad examina la estabilidad del procedimiento. Li et al. (2024) demostraron que el rendimiento de los modelos puede variar al aplicar criterios de clasificación documental. Por ello, una respuesta técnicamente correcta en una ejecución no garantiza resultados consistentes ni científicamente válidos en otros contextos.

La validez interna permite determinar si los resultados se derivan del procedimiento aplicado y no de factores externos, errores de medición o decisiones inadecuadas. En investigaciones asistidas por IA, esta evaluación comprende la calidad de los datos, la correspondencia entre variables, la selección del modelo y el control de posibles factores de confusión. Un sistema puede producir resultados coherentes y, sin embargo, apoyarse en relaciones irrelevantes. La revisión debe establecer si el modelo aprendió características vinculadas con el fenómeno o señales accidentales presentes únicamente en el conjunto analizado.

La validez externa examina hasta qué punto los resultados pueden trasladarse hacia otras poblaciones, periodos, instituciones o condiciones experimentales. El rendimiento obtenido con una muestra específica no garantiza que el modelo conserve su utilidad cuando cambian la procedencia de los datos o las características del contexto. Barberis et al. (2024) destacan la necesidad de fortalecer la robustez y reproducibilidad de los modelos utilizados en ciencias biomédicas. La generalización debe comprobarse mediante datos independientes y escenarios suficientemente distintos de aquellos empleados durante el desarrollo inicial.

La precisión representa la proporción de resultados correctamente clasificados, pero no debe confundirse con validez científica. Un modelo puede alcanzar una precisión elevada si predomina una categoría y aun así presentar dificultades para reconocer los casos menos frecuentes. Por ello, conviene complementar esta medida con sensibilidad y especificidad. La sensibilidad indica la capacidad para identificar correctamente los casos positivos, mientras la especificidad evalúa el reconocimiento de los negativos. La

selección de indicadores debe responder a las consecuencias de cada error y no únicamente al valor más favorable.

Syriani et al. (2024) muestran que los modelos generativos pueden apoyar el cribado de publicaciones, aunque su desempeño depende de los criterios y las instrucciones utilizadas. En este tipo de tareas, un falso negativo puede excluir evidencia relevante, mientras un falso positivo aumenta la carga de revisión. La evaluación debe considerar ambos tipos de error y determinar cuál resulta más perjudicial para el propósito científico. Los indicadores necesitan interpretarse según la tarea, porque una misma puntuación puede ser aceptable en exploración preliminar y resultar insuficiente para decisiones definitivas.

La confiabilidad también comprende estabilidad entre ejecuciones. Los modelos generativos pueden ofrecer respuestas diferentes ante una misma instrucción debido a su naturaleza probabilística, cambios de versión o configuraciones internas. Behrend y Landers (2025) señalan que su uso en materiales de investigación requiere procedimientos de revisión y control. Para evaluar estabilidad, conviene repetir solicitudes, comparar salidas y documentar parámetros. La variabilidad no invalida automáticamente el procedimiento, pero debe cuantificarse y comunicarse cuando afecta la clasificación, la interpretación o las conclusiones producidas.

La comparación con expertos permite valorar si el sistema alcanza resultados semejantes a los obtenidos mediante evaluación humana. Sin embargo, el juicio experto tampoco constituye una referencia perfecta, porque puede existir desacuerdo entre evaluadores. Girdwood y Manikonda (2026) plantean que la incorporación de IA en investigación requiere decisiones prácticas y metodológicas durante todo el proceso. Una comparación rigurosa debe emplear criterios explícitos, varios evaluadores y mecanismos para resolver discrepancias, evitando presentar la coincidencia con una sola persona como prueba suficiente de confiabilidad.

La validación cruzada ayuda a evaluar el comportamiento del modelo mediante divisiones sucesivas de los datos disponibles. Este procedimiento permite estimar si el rendimiento depende excesivamente de una partición particular. Sin embargo, la validación cruzada permanece dentro del mismo universo de datos y puede conservar sesgos, errores o particularidades del conjunto original. Por ello, debe complementarse con validación externa. Los resultados son más sólidos cuando se mantienen en muestras independientes,

recogidas mediante otros procedimientos y representativas de condiciones diferentes a las utilizadas para entrenar o ajustar el sistema.

La robustez indica si el modelo conserva un funcionamiento razonable cuando existen pequeñas variaciones, ruido o datos incompletos. Un sistema frágil puede modificar significativamente sus resultados ante cambios mínimos en la entrada. Messeri y Crockett (2024) advierten que la inteligencia artificial puede generar ilusiones de comprensión cuando el rendimiento se interpreta como explicación suficiente. Evaluar robustez requiere introducir perturbaciones controladas, comparar configuraciones y analizar casos adversos. También implica identificar las condiciones específicas bajo las cuales el modelo deja de responder de manera confiable.

En consecuencia, la calidad de los resultados asistidos por inteligencia artificial debe evaluarse mediante varios criterios complementarios. Precisión, sensibilidad y especificidad describen aspectos del rendimiento, pero no sustituyen la valoración de validez interna, generalización, estabilidad y robustez. La confiabilidad científica aumenta cuando los resultados se comparan con expertos, métodos alternativos y datos externos. Ningún indicador aislado demuestra que un sistema sea válido para todas las condiciones. Las conclusiones deben reconocer incertidumbres, variabilidad y límites de aplicación, manteniendo correspondencia entre evidencia, procedimiento y alcance declarado.

Tabla 13
Criterios para evaluar la validez y confiabilidad de resultados asistidos por inteligencia artificial

Criterio de calidad	Pregunta de evaluación	Indicador posible	Evidencia requerida
Validez interna	¿El resultado se deriva realmente de las variables y procedimientos definidos?	Coherencia entre diseño, datos, modelo y conclusión	Control de factores de confusión, revisión metodológica y análisis de errores
Validez externa	¿El resultado puede mantenerse en otras poblaciones o contextos?	Rendimiento en muestras independientes	Datos externos procedentes de otros periodos, instituciones o poblaciones
Precisión	¿Qué proporción de resultados fue	Porcentaje global de aciertos	Matriz de confusión y comparación con resultados conocidos

Criterio de calidad	Pregunta de evaluación	Indicador posible	Evidencia requerida
	clasificada correctamente?		
Sensibilidad	¿El sistema reconoce correctamente los casos positivos?	Tasa de verdaderos positivos	Conjunto de casos positivos confirmado mediante referencia independiente
Especificidad	¿El sistema reconoce correctamente los casos negativos?	Tasa de verdaderos negativos	Conjunto de casos negativos comprobado y criterios claros de exclusión
Estabilidad	¿El sistema produce resultados semejantes en ejecuciones comparables?	Variación entre repeticiones	Ejecuciones múltiples con instrucciones, datos y parámetros documentados
Validación cruzada	¿El rendimiento se mantiene en diferentes particiones de los datos?	Promedio y dispersión entre particiones	División sistemática del conjunto y repetición del entrenamiento y evaluación
Generalización	¿El modelo funciona con datos no utilizados durante su desarrollo?	Diferencia entre rendimiento interno y externo	Conjunto independiente sin participación en entrenamiento o ajuste
Robustez	¿El resultado se mantiene ante ruido, datos incompletos o pequeñas variaciones?	Cambio de rendimiento bajo perturbaciones	Pruebas adversariales, análisis de sensibilidad y escenarios modificados
Comparación con expertos	¿El sistema coincide con evaluaciones humanas fundamentadas?	Concordancia entre sistema y especialistas	Varios expertos, criterios comunes y procedimiento para resolver discrepancias
Variabilidad	¿Cuánto cambian las salidas entre modelos, versiones o configuraciones?	Dispersión y consistencia de resultados	Comparaciones entre versiones, parámetros y sistemas alternativos
Utilidad científica	¿El resultado responde de manera válida a la pregunta de investigación?	Correspondencia entre evidencia y objetivo	Interpretación disciplinar, revisión metodológica y contraste con conocimiento previo

Nota. La tabla organiza criterios complementarios para evaluar resultados asistidos por inteligencia artificial. Un rendimiento elevado no demuestra por sí solo validez científica; se requiere evidencia sobre estabilidad, generalización, robustez y correspondencia con el problema investigado.

La validez y confiabilidad de los resultados asistidos por inteligencia artificial deben evaluarse mediante criterios complementarios y no a partir de una sola métrica. La

precisión, la sensibilidad y la especificidad describen el rendimiento, pero no garantizan generalización, estabilidad ni utilidad científica. Por ello, los resultados deben compararse con datos externos, expertos, modelos alternativos y condiciones modificadas. Solo mediante esta evaluación integral es posible determinar si una salida algorítmica responde de manera consistente, robusta y metodológicamente válida a la pregunta de investigación.

4.2 Reproducibilidad y trazabilidad computacional

La reproducibilidad permite que otros investigadores repitan un procedimiento utilizando los mismos datos, código y condiciones documentadas, mientras que la trazabilidad permite reconstruir cómo se obtuvo cada resultado. En investigaciones asistidas por inteligencia artificial, ambas exigencias adquieren mayor complejidad porque intervienen modelos, versiones, parámetros, instrucciones y servicios externos. Masbernat (2024) vincula la integridad científica con la posibilidad de verificar los pasos realizados. Por ello, documentar el proceso computacional no constituye una tarea administrativa, sino una condición para evaluar la confiabilidad del conocimiento producido.

El registro de los datos debe identificar su procedencia, fecha de obtención, formato, condiciones de acceso y transformaciones realizadas. También conviene conservar una copia inalterada del conjunto original y diferenciarla de las versiones utilizadas en limpieza, análisis o entrenamiento. Choque et al. (2025) relacionan la ciencia abierta con la disponibilidad y reutilización responsable de recursos científicos. Sin esta separación, resulta difícil determinar si una diferencia en los resultados proviene del modelo, del código o de modificaciones efectuadas sobre la información inicial.

El código empleado debe conservarse junto con instrucciones suficientes para ejecutarlo. No basta con almacenar el archivo final, porque los análisis suelen desarrollarse mediante correcciones, pruebas y cambios sucesivos. La gestión de versiones permite registrar qué líneas fueron modificadas, quién realizó el cambio y cuál fue su propósito. Cornelio et al. (2023) muestran la importancia de integrar procedimientos computacionales con condiciones teóricas explícitas. Una reproducción rigurosa requiere conocer tanto la secuencia técnica como los supuestos científicos incorporados en cada etapa del análisis.

Las versiones de modelos, bibliotecas y servicios también deben quedar documentadas. Un mismo nombre comercial puede corresponder a sistemas actualizados con

comportamientos distintos, y una biblioteca puede modificar funciones entre versiones. Por ello, conviene registrar identificadores, fechas, dependencias y configuraciones utilizadas. Van et al. (2024) proponen mediante ENCORE una organización práctica para conservar datos, código, resultados y entornos computacionales. Esta documentación permite distinguir cambios metodológicos de variaciones producidas por actualizaciones tecnológicas ajenas al diseño original del estudio.

Los parámetros determinan cómo funciona un modelo y pueden modificar sustancialmente sus resultados. Deben registrarse configuraciones de entrenamiento, criterios de optimización, semillas aleatorias, temperatura, límites de longitud y cualquier ajuste aplicado durante la ejecución. Barberis et al. (2024) destacan que la robustez y reproducibilidad requieren procedimientos estandarizados y comparables. Omitir parámetros impide saber si una diferencia responde al azar, a una configuración distinta o a un error. La transparencia exige documentar incluso aquellas decisiones consideradas secundarias durante el desarrollo inicial.

En los modelos generativos, las instrucciones o prompts forman parte del procedimiento metodológico. Debe conservarse el texto exacto, el contexto proporcionado, el orden de las interacciones y los documentos incorporados. También resulta necesario registrar modificaciones posteriores, porque pequeñas variaciones pueden producir respuestas diferentes. Shao et al. (2025) muestran que los sistemas colaborativos pueden organizar flujos analíticos complejos, pero su repetición depende de conservar las condiciones de interacción. El prompt no es una nota informal; constituye un insumo que debe tratarse como parte del protocolo.

La fecha y el entorno de ejecución ayudan a reconstruir las condiciones técnicas bajo las cuales se generaron los resultados. Deben incluirse sistema operativo, capacidad de cómputo, bibliotecas, dispositivos, servicios utilizados y posibles restricciones de acceso. Cuando la investigación depende de plataformas externas, también conviene señalar si el modelo era local, remoto o accesible mediante una interfaz. Estas condiciones pueden influir en velocidad, precisión, disponibilidad y formato de las salidas. Registrar el entorno permite explicar diferencias que no provienen directamente del razonamiento científico.

Las transformaciones de datos deben describirse de manera secuencial y verificable. Esto incluye eliminación de duplicados, normalización, imputación, etiquetado,

anonimización y conversión de formatos. Cada modificación debe asociarse con una regla y una justificación. Van et al. (2024) resaltan la utilidad de estructuras que conectan entradas, procedimientos y resultados. La trazabilidad se pierde cuando una base final aparece sin explicación sobre cómo fue construida. En esas condiciones, otros investigadores pueden repetir el código, pero no reconstruir el conjunto que realmente produjo las conclusiones.

Las decisiones metodológicas también requieren documentación. Deben registrarse modelos descartados, criterios de selección, errores encontrados, cambios de estrategia y resultados que no fueron incorporados al manuscrito. Esta información permite comprender por qué se eligió una ruta y no otra. Masbernat (2024) relaciona la integridad científica con transparencia y mantenimiento de registros. La trazabilidad no debe limitarse a conservar aquello que funcionó; también debe mostrar alternativas rechazadas, porque esas decisiones influyen en la interpretación y pueden revelar sesgos durante el proceso.

Los resultados intermedios permiten comprobar si cada etapa funcionó según lo previsto. Conviene conservar salidas de limpieza, métricas preliminares, visualizaciones de diagnóstico y registros de ejecución. Shao et al. (2025) plantean que los flujos asistidos por inteligencia artificial pueden favorecer la iteración y la reproducibilidad cuando sus pasos permanecen organizados. Si solo se almacena el resultado final, un error temprano puede quedar oculto. Los productos intermedios facilitan auditorías, comparación entre versiones y detección del punto exacto donde surgió una discrepancia.

La reproducibilidad no implica necesariamente obtener una respuesta textual idéntica, especialmente cuando se utilizan modelos probabilísticos. Lo importante es que otros investigadores puedan repetir el procedimiento, comprender sus condiciones y obtener resultados suficientemente comparables. Los repositorios deben reunir datos, código, documentación, versiones y registros de decisiones con permisos adecuados. La auditoría permite revisar cómo se generó cada conclusión y qué controles fueron aplicados. Una investigación trazable no elimina la variabilidad, pero hace posible explicarla, evaluarla y determinar si afecta el alcance científico declarado.

Tabla 14

Lista de verificación para la reproducibilidad y trazabilidad computacional

Elemento	Información que debe registrarse	Formato recomendado	Finalidad
Datos originales	Fuente, fecha, identificador, formato, licencia y condiciones de acceso	Repositorio con versión inalterada y metadatos	Conservar el punto de partida del estudio
Datos procesados	Limpieza, exclusiones, normalización, imputación y nuevas variables	Archivos versionados y bitácora de transformaciones	Reconstruir cómo se obtuvo la base analítica
Código	Scripts, funciones, dependencias y orden de ejecución	Repositorio con control de versiones	Permitir repetición, revisión y corrección
Modelo utilizado	Nombre, versión, proveedor, fecha y forma de acceso	Ficha técnica o archivo de configuración	Identificar el sistema exacto empleado
Parámetros	Semillas, temperatura, tasas, umbrales y configuraciones	Archivo estructurado de parámetros	Explicar variaciones entre ejecuciones
Prompts	Texto exacto, contexto, archivos suministrados y secuencia	Registro textual o archivo de sesión	Repetir la interacción con modelos generativos
Entorno computacional	Sistema operativo, hardware, bibliotecas y servicios	Archivo de entorno o contenedor	Reproducir las condiciones técnicas
Fechas de ejecución	Inicio, finalización, actualizaciones y repeticiones	Bitácora cronológica	Relacionar resultados con versiones y condiciones temporales
Decisiones metodológicas	Modelos probados, criterios, cambios y alternativas descartadas	Diario de investigación o registro de decisiones	Justificar la ruta analítica seleccionada
Resultados intermedios	Métricas, diagnósticos, visualizaciones y errores	Carpetas versionadas y registros automáticos	Detectar dónde surgieron diferencias o fallas
Resultados finales	Tablas, figuras, modelos y archivos utilizados en la publicación	Repositorio con identificadores persistentes	Vincular el manuscrito con la evidencia producida
Auditoría	Responsables, revisiones, observaciones y correcciones	Informe de auditoría o lista de comprobación	Verificar integridad, coherencia y trazabilidad

Nota. La tabla presenta los elementos mínimos que deben documentarse para reconstruir una investigación computacional asistida por inteligencia artificial. La reproducibilidad depende de conservar datos, código, modelos, parámetros, instrucciones y decisiones metodológicas dentro de un entorno organizado y auditable.

La reproducibilidad y la trazabilidad computacional requieren documentar de forma organizada cada componente del proceso investigativo, desde los datos originales hasta los resultados finales. El registro de versiones, parámetros, prompts, entornos de ejecución y decisiones metodológicas permite reconstruir el análisis, explicar variaciones y localizar posibles errores. Esta documentación no solo facilita la repetición técnica del estudio, sino también su auditoría y evaluación crítica. Una investigación asistida por inteligencia artificial resulta más confiable cuando sus procedimientos, transformaciones y responsabilidades pueden verificarse de manera clara.

4.3 Explicabilidad e interpretabilidad de los modelos

La explicabilidad y la interpretabilidad son condiciones relevantes para evaluar cómo un modelo produce resultados y hasta qué punto estos pueden comprenderse dentro de una investigación científica. Aunque ambos conceptos suelen utilizarse como equivalentes, responden a preguntas distintas. La interpretabilidad se relaciona con la posibilidad de comprender la relación entre entradas, variables y resultados, mientras la explicabilidad busca ofrecer razones inteligibles sobre una decisión específica. Chaaban et al. (2024) vinculan estas capacidades con la preparación crítica necesaria para utilizar inteligencia artificial de manera responsable en investigación.

La transparencia se refiere al grado en que pueden conocerse la arquitectura, los datos, las reglas y los procedimientos empleados por el sistema. Un modelo transparente permite identificar sus componentes y reconstruir las operaciones principales que conducen a una salida. Sin embargo, disponer de información técnica no garantiza que el investigador comprenda realmente el resultado. Un sistema puede estar completamente documentado y continuar siendo difícil de interpretar debido a su complejidad matemática. Por ello, la transparencia constituye una condición favorable, pero no suficiente para lograr comprensión científica.

La interpretabilidad se centra en reconocer cómo determinadas variables o características influyen en el resultado producido. En modelos simples, esta relación puede observarse directamente mediante coeficientes, reglas o estructuras visibles. En sistemas complejos, la interpretación suele apoyarse en métodos adicionales que estiman la importancia de variables o representan relaciones internas. Behrend y Landers (2025) señalan que los investigadores deben comprender el funcionamiento general de los modelos utilizados,

especialmente cuando estos intervienen en la generación de materiales, decisiones o resultados que afectan el diseño del estudio.

La explicabilidad busca traducir el comportamiento del modelo en una justificación comprensible para investigadores, participantes o usuarios. Una explicación puede indicar qué variables fueron determinantes, cómo cambió la predicción o qué condiciones influyeron en una clasificación. Schwartz (2022) plantea que no todos los sistemas necesitan el mismo grado de interpretabilidad, pues la exigencia depende del propósito y de las consecuencias del resultado. En investigación científica, las explicaciones deben permitir evaluar coherencia, límites y posibles errores, no limitarse a ofrecer una narrativa técnicamente convincente.

La justificación científica posee un alcance más amplio que la explicación técnica. Un modelo puede explicar qué variables influyeron en su salida sin demostrar que esa relación sea válida, causal o relevante para el fenómeno estudiado. La justificación requiere vincular el resultado con datos, teoría, diseño metodológico y evidencia independiente. Por tanto, conocer cómo se produjo una predicción no equivale a saber por qué ocurre el fenómeno. Esta diferencia evita que una descripción del comportamiento algorítmico sea presentada como explicación sustantiva de la realidad científica investigada.

Los modelos de caja negra presentan dificultades porque sus procesos internos no pueden reconstruirse fácilmente, aun cuando ofrecen resultados precisos. Esta situación es común en sistemas con numerosas capas, parámetros e interacciones. Krenn et al. (2022) sostienen que el valor científico de la inteligencia artificial debe evaluarse por su capacidad para favorecer comprensión y no solamente por su rendimiento predictivo. Un modelo puede ser útil para anticipar resultados, pero insuficiente cuando la investigación busca descubrir mecanismos, formular teorías o justificar decisiones con consecuencias relevantes.

Las explicaciones locales describen por qué el modelo produjo un resultado específico para una observación determinada. Pueden indicar qué variables influyeron en una clasificación individual o cómo pequeñas modificaciones cambiarían la salida. Las explicaciones globales intentan representar el comportamiento general del sistema, identificando patrones, reglas o relaciones predominantes. Ambos niveles son complementarios. Una explicación local puede ser clara, pero no describir el modelo

completo; una explicación global puede mostrar tendencias generales, aunque resultar insuficiente para comprender casos particulares o excepciones importantes.

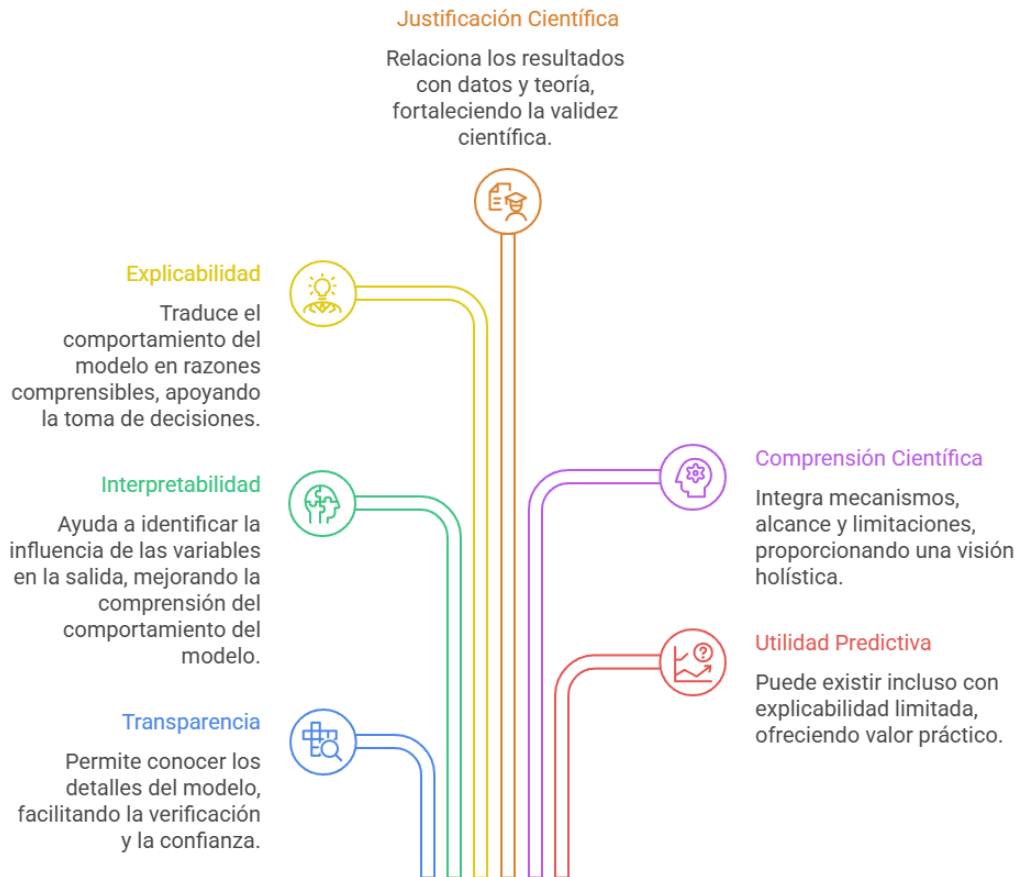
La importancia de variables constituye una estrategia frecuente para interpretar modelos, pero debe analizarse con prudencia. Una variable puede aparecer como influyente porque está correlacionada con otra, porque refleja un sesgo del conjunto o porque el modelo la utiliza como indicador indirecto. Su relevancia algorítmica no demuestra importancia causal ni significado disciplinar. El investigador debe contrastar estas estimaciones con conocimiento teórico, análisis de sensibilidad y datos externos. De lo contrario, podría construir explicaciones científicas sobre relaciones que solo describen el funcionamiento interno del sistema.

Las explicaciones engañosas representan un riesgo adicional, porque pueden parecer claras y plausibles sin reflejar fielmente el proceso del modelo. Messeri y Crockett (2024) advierten que la inteligencia artificial puede generar ilusiones de comprensión cuando ofrece respuestas organizadas que reducen la percepción de incertidumbre. Una explicación debe evaluarse por su fidelidad, estabilidad y capacidad para anticipar cambios en la salida. También conviene comparar distintos métodos explicativos, pues resultados contradictorios indican que la interpretación depende de supuestos o aproximaciones que necesitan hacerse explícitos.

En consecuencia, transparencia, interpretabilidad, explicabilidad, justificación y comprensión científica deben considerarse niveles relacionados, pero no intercambiables. La transparencia muestra qué sistema fue utilizado; la interpretabilidad ayuda a reconocer relaciones internas; la explicabilidad comunica razones comprensibles; la justificación conecta el resultado con evidencia y método; y la comprensión científica integra mecanismos, teoría y alcance. Un modelo puede ser predictivamente útil sin cumplir todos estos niveles, pero cuanto mayor sea la consecuencia de sus resultados, mayor debe ser la exigencia de evaluación, documentación y revisión humana.

Figura 12

Relación entre transparencia, interpretabilidad, explicabilidad y comprensión científica



Nota. La figura diferencia niveles relacionados con la comprensión de modelos de inteligencia artificial. La utilidad predictiva no garantiza por sí sola explicabilidad, justificación ni comprensión científica; estas requieren evidencia adicional, análisis metodológico e interpretación disciplinar.

La transparencia, la interpretabilidad y la explicabilidad permiten examinar distintos niveles del funcionamiento de un modelo, pero no equivalen automáticamente a comprensión científica. Conocer los datos, parámetros o variables influyentes ayuda a evaluar una salida, aunque la justificación exige relacionarla con teoría, diseño metodológico y evidencia independiente. Del mismo modo, una predicción útil puede carecer de explicación suficiente. Por ello, el valor científico de un modelo depende de su capacidad para apoyar interpretaciones verificables, reconocer limitaciones y contribuir a una comprensión disciplinar fundamentada.

4.4 Sesgos algorítmicos y representatividad de los datos

Los sesgos algorítmicos surgen cuando los datos, las decisiones de diseño o los procedimientos de evaluación favorecen sistemáticamente determinadas representaciones y perjudican otras. En investigación científica, este problema no se limita a la

discriminación social, porque también puede afectar la selección de variables, la cobertura geográfica, la diversidad disciplinar y la interpretación de resultados. Chaaban et al. (2024) destacan que la preparación para utilizar inteligencia artificial requiere reconocer sus limitaciones éticas y contextuales. Un modelo puede ser técnicamente preciso y, al mismo tiempo, producir conocimiento parcial o poco generalizable.

Los sesgos históricos aparecen cuando los datos reproducen desigualdades, decisiones o patrones acumulados en periodos anteriores. Si un sistema aprende de registros contruidos bajo condiciones inequitativas, puede tratar esas relaciones como si representaran el funcionamiento normal del fenómeno. Masbernat (2024) advierte que la integridad científica exige examinar la calidad y procedencia de la información utilizada. El modelo no crea necesariamente el sesgo desde cero, pero puede amplificarlo, estabilizarlo y presentarlo como una regularidad objetiva cuando en realidad refleja prácticas sociales o institucionales previas.

El sesgo de selección se produce cuando las observaciones incluidas no representan adecuadamente la población o el universo sobre el cual se formularán conclusiones. Puede originarse por acceso limitado, participación voluntaria, exclusión de casos difíciles o dependencia de fuentes digitalizadas. Truhn et al. (2024) muestran que los modelos multimodales requieren datos diversos para generar resultados aplicables en contextos clínicos. Si determinados grupos aparecen en proporciones reducidas, el sistema puede funcionar correctamente para la mayoría y presentar errores elevados precisamente en las poblaciones menos representadas.

El etiquetado introduce otra fuente de sesgo porque las categorías suelen depender de decisiones humanas. Diferentes especialistas pueden interpretar de manera distinta una imagen, un discurso o una conducta, y esas diferencias pasan a formar parte del conjunto utilizado para entrenar el modelo. Las etiquetas también pueden simplificar fenómenos complejos mediante categorías rígidas. Por ello, deben documentarse los criterios empleados, la formación de los evaluadores y el nivel de concordancia alcanzado. Una clasificación automática reproduce las definiciones incorporadas, incluso cuando estas resultan ambiguas o culturalmente limitadas.

El idioma y la cultura condicionan la representatividad de los sistemas generativos y analíticos. Los modelos entrenados principalmente con materiales en inglés o procedentes de contextos dominantes pueden comprender peor variedades lingüísticas, expresiones

regionales y conocimientos producidos en otras tradiciones. Messeri y Crockett (2024) señalan que la dependencia de herramientas homogéneas puede generar monoculturas científicas e ilusiones de comprensión. Cuando los mismos modelos median búsquedas, análisis y redacción, ciertos conceptos y estilos de razonamiento pueden adquirir mayor visibilidad, mientras otros permanecen subrepresentados o son interpretados mediante categorías externas.

La geografía también influye en la composición de los datos y de la literatura científica disponible. Regiones con mayor infraestructura, financiamiento y presencia en bases internacionales producen más registros accesibles para los sistemas. En consecuencia, una revisión o modelo puede identificar como universales patrones derivados principalmente de países con alta producción digital. Choque et al. (2025) relacionan la ciencia abierta con la necesidad de ampliar el acceso y la colaboración. Sin embargo, la apertura técnica no resuelve automáticamente desigualdades de producción, idioma, conectividad y capacidad institucional.

El sesgo disciplinar aparece cuando los modelos privilegian métodos, conceptos o criterios procedentes de áreas con mayor disponibilidad de datos. Un sistema entrenado con abundante literatura biomédica o computacional puede trasladar esas formas de evaluación hacia ciencias sociales, humanidades o investigación cualitativa. Esto favorece la homogeneización metodológica y reduce la diversidad epistemológica. La inteligencia artificial debe adaptarse a las exigencias de cada campo, porque los criterios de evidencia, validez e interpretación no son idénticos. Tang et al. (2025) advierten que la autonomía científica requiere salvaguardas frente a decisiones amplificadas sin supervisión suficiente.

Los sesgos de publicación afectan la información que los modelos utilizan para aprender o sintetizar conocimiento. Los resultados positivos, novedosos o estadísticamente significativos suelen publicarse con mayor frecuencia que los hallazgos nulos o las replicaciones. Si la inteligencia artificial procesa principalmente literatura publicada, puede sobreestimar la consistencia de determinados efectos y presentar consensos más sólidos de lo que realmente son. También puede ignorar datos negativos, informes técnicos y conocimientos locales. La representatividad documental exige buscar fuentes diversas y reconocer qué tipos de evidencia quedaron fuera del corpus.

La evaluación por grupos permite detectar si el rendimiento cambia según población, idioma, región, disciplina o fuente de datos. No basta con calcular una medida global, porque un promedio favorable puede ocultar errores concentrados en subgrupos pequeños. Deben compararse sensibilidad, precisión, tasas de error y estabilidad en cada segmento relevante. También conviene realizar análisis de intersección, ya que varias características pueden combinarse y producir desventajas específicas. La evaluación debe acompañarse de una explicación sobre por qué existen diferencias y qué consecuencias generan para el alcance de las conclusiones.

La mitigación de sesgos requiere intervenir en datos, modelos, evaluación y comunicación. Puede incluir ampliar muestras, equilibrar categorías, revisar etiquetas, incorporar fuentes regionales, comparar modelos y ajustar criterios de decisión. Sin embargo, ninguna técnica elimina completamente el problema. La documentación debe indicar qué grupos están representados, cuáles faltan y bajo qué condiciones se evaluó el sistema. La transparencia permite que otros investigadores determinen si los resultados son aplicables a sus contextos. Mitigar no significa ocultar diferencias, sino reconocerlas y reducir sus efectos injustificados.

En consecuencia, la representatividad debe considerarse una propiedad científica que afecta validez, generalización e interpretación. Un sistema puede procesar grandes cantidades de datos y continuar describiendo solo una parte limitada de la realidad. Los sesgos históricos, lingüísticos, geográficos, disciplinares y de publicación muestran que la cantidad no sustituye la diversidad ni la calidad. La evaluación responsable exige identificar quiénes, qué conocimientos y qué contextos están presentes en los datos, además de comunicar con precisión dónde terminan las posibilidades de generalización del modelo.

Tabla 15

Tipos de sesgos algorítmicos y estrategias para su evaluación

Tipo de sesgo	Origen principal	Efecto científico	Forma de detección	Estrategia de mitigación
Histórico	Registros producidos bajo desigualdades o prácticas anteriores	Reproducción de relaciones injustas como si fueran normales	Análisis temporal y contextual de los datos	Revisar variables, incorporar conocimiento histórico y limitar conclusiones

Tipo de sesgo	Origen principal	Efecto científico	Forma de detección	Estrategia de mitigación
Selección	Muestras incompletas o acceso desigual a participantes y fuentes	Generalización inadecuada hacia poblaciones no representadas	Comparación entre muestra, población objetivo y fuentes excluidas	Ampliar cobertura, utilizar muestreo complementario y ponderar resultados
Etiquetado	Criterios humanos ambiguos, inconsistentes o culturalmente restringidos	Clasificaciones inestables o simplificadas	Concordancia entre evaluadores y revisión de casos dudosos	Definir protocolos, capacitar evaluadores y conservar múltiples perspectivas
Desbalance de datos	Predominio de una categoría o grupo	Buen rendimiento global y errores elevados en minorías	Métricas separadas por clase y matriz de confusión	Recolección adicional, remuestreo y ajuste de criterios
Lingüístico	Predominio de determinados idiomas y variedades	Menor comprensión de expresiones locales o regionales	Evaluación por idioma, dialecto y registro	Incorporar corpus multilingües y validación con hablantes locales
Cultural	Uso de categorías ajenas al contexto estudiado	Interpretaciones descontextualizadas o estereotipadas	Revisión por especialistas y participantes del contexto	Adaptación cultural, análisis cualitativo y consulta comunitaria
Geográfico	Concentración de datos en países o regiones con mayor infraestructura	Presentación de patrones locales como universales	Mapas de procedencia y comparación regional	Incluir fuentes regionales y limitar el alcance declarado
Disciplinar	Predominio de métodos y conceptos de campos con mayor digitalización	Homogeneización de criterios científicos	Comparación entre marcos y prácticas disciplinares	Diseñar modelos y evaluaciones específicos para cada campo
De publicación	Mayor presencia de resultados positivos o novedosos	Sobreestimación de efectos y consensos	Revisión de literatura gris, registros y resultados nulos	Incorporar repositorios, preprints y estudios no significativos
De medición	Instrumentos que capturan de forma desigual el fenómeno	Variables incompletas o comparaciones inválidas	Evaluación de validez y consistencia de instrumentos	Ajustar medidas, calibrar equipos y combinar fuentes

Tipo de sesgo	Origen principal	Efecto científico	Forma de detección	Estrategia de mitigación
Algorítmico	Funciones objetivo, umbrales o procedimientos de optimización	Priorización sistemática de ciertos resultados	Análisis de sensibilidad y comparación entre modelos	Modificar objetivos, umbrales y restricciones
De evaluación	Uso de métricas globales que ocultan diferencias	Percepción errónea de rendimiento uniforme	Indicadores desagregados por grupos y contextos	Evaluar subgrupos, intersecciones y condiciones externas

Nota. La tabla muestra que los sesgos pueden originarse en los datos, el diseño, el contexto y los criterios de evaluación. Su mitigación requiere acciones combinadas y una documentación explícita sobre cobertura, limitaciones y alcance de los resultados.

Los sesgos algorítmicos pueden originarse en los datos, las categorías, los modelos, los contextos y las métricas utilizadas para evaluar el rendimiento. Sus efectos no se limitan a la discriminación social, sino que también pueden reducir la validez, la representatividad y la capacidad de generalización de los resultados científicos. Por ello, su detección exige análisis desagregados, revisión contextual y comparación entre grupos, regiones y disciplinas. La mitigación requiere combinar mejores datos, ajustes metodológicos, validación externa y documentación transparente de las limitaciones.

4.5 Alucinaciones, errores y verificación de resultados generativos

Las alucinaciones generativas son respuestas que presentan información falsa, imprecisa o no sustentada con una forma lingüística coherente y convincente. En investigación científica, pueden manifestarse como datos inventados, referencias inexistentes, explicaciones conceptuales incorrectas o resultados atribuidos indebidamente a una fuente. Athaluri et al. (2023) comprobaron este problema al examinar referencias producidas por ChatGPT en propuestas científicas. La fluidez textual aumenta el riesgo de aceptación, porque una afirmación bien redactada puede parecer confiable aunque no exista evidencia que permita confirmarla.

Las referencias inexistentes constituyen uno de los errores más reconocibles. El sistema puede combinar nombres reales, títulos plausibles, revistas conocidas y datos bibliográficos inventados, creando una cita aparentemente completa. También puede proporcionar un DOI existente que conduce a otra publicación o atribuir a un artículo real conclusiones que este no contiene. Por ello, no basta con comprobar que el documento

existe. Cada referencia debe verificarse en la página editorial, el repositorio institucional o una base bibliográfica confiable, contrastando autoría, título, año, DOI y contenido efectivamente respaldado.

Los datos fabricados incluyen porcentajes, tamaños muestrales, fechas, coeficientes o resultados experimentales que no proceden de una fuente identificable. Estos valores pueden incorporarse a tablas o argumentos y adquirir apariencia de evidencia empírica. Lopezosa (2023) sostiene que la utilización académica de sistemas generativos debe acompañarse de revisión crítica y comprobación documental. La verificación requiere localizar la fuente primaria, revisar el apartado donde aparece la cifra y confirmar su unidad, población, periodo y contexto. Cuando no existe respaldo directo, el dato debe eliminarse, incluso si parece razonable.

También pueden aparecer errores conceptuales cuando el modelo mezcla definiciones, confunde relaciones o aplica términos de una disciplina en un sentido inadecuado. Messeri y Crockett (2024) advierten que la inteligencia artificial puede generar una ilusión de comprensión al ofrecer explicaciones organizadas que ocultan incertidumbres y simplifican problemas complejos. Para verificar conceptos, conviene compararlos con artículos originales, textos especializados y marcos teóricos reconocidos. La revisión debe determinar si la explicación describe realmente el fenómeno, si conserva sus condiciones y si distingue correlación, predicción, causalidad e interpretación.

Los errores matemáticos y estadísticos pueden surgir en operaciones, fórmulas, interpretación de indicadores o selección de procedimientos. Del mismo modo, el código generado puede ejecutarse sin interrupciones y producir resultados incorrectos debido a variables mal definidas, funciones inadecuadas o transformaciones silenciosas. La comprobación debe incluir cálculos manuales sobre casos simples, pruebas con datos conocidos y comparación con otro programa o procedimiento independiente. Además, es necesario inspeccionar las entradas, salidas intermedias y supuestos estadísticos, porque una respuesta numéricamente plausible no demuestra que el razonamiento aplicado sea válido.

La información desactualizada representa otro riesgo, especialmente cuando los modelos no identifican con claridad la fecha límite de sus datos o combinan versiones diferentes de una norma, herramienta o evidencia científica. Zhang et al. (2025) explican que los grandes modelos de lenguaje pueden apoyar distintas fases del método científico, pero

sus aportes necesitan evaluación y validación. Toda afirmación temporal debe contrastarse con fuentes recientes, indicando fecha de consulta y versión. Este control es esencial para legislación, guías clínicas, programas informáticos, estadísticas oficiales y áreas científicas que cambian rápidamente.

El exceso de confianza aparece cuando el sistema expresa una respuesta incierta mediante afirmaciones categóricas, sin reconocer alternativas ni límites. Goyanes y Lopezosa (2024) señalan que la colaboración humano-máquina requiere pensamiento crítico y supervisión constante. Para detectar este problema, debe solicitarse al modelo que separe hechos, inferencias y supuestos, pero sus propias declaraciones de confianza no deben considerarse prueba. La seguridad de una frase no equivale a exactitud. El investigador necesita buscar evidencia externa, identificar desacuerdos y comunicar el grado de certeza que realmente permiten los datos consultados.

La triangulación consiste en contrastar una afirmación mediante fuentes, métodos o evaluadores independientes. Puede incluir la consulta de dos publicaciones primarias, la repetición del cálculo con otro programa y la revisión por un especialista. Sin embargo, consultar varios modelos generativos no constituye triangulación suficiente si todos dependen de patrones o materiales semejantes. La independencia de la evidencia resulta fundamental. Tang et al. (2025) recomiendan priorizar salvaguardas y controles frente a una autonomía científica creciente, debido a que los errores pueden propagarse entre tareas conectadas y producir consecuencias acumulativas.

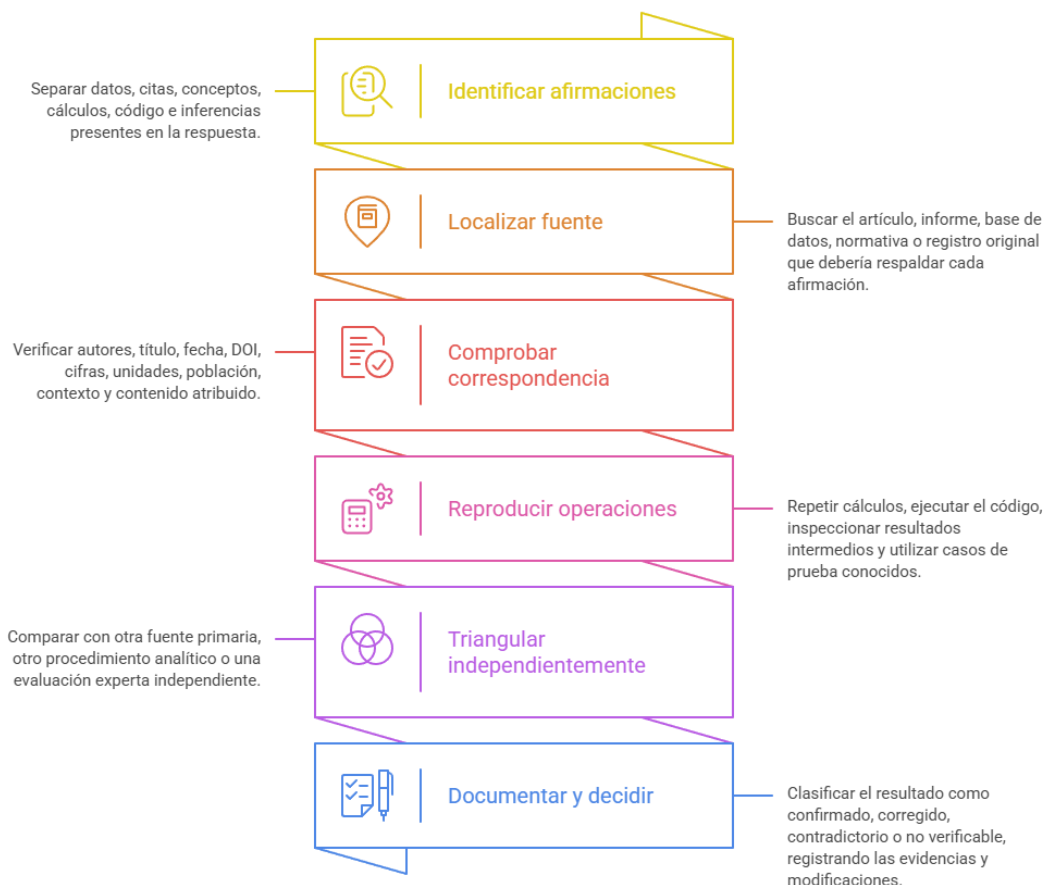
La verificación debe concentrarse primero en afirmaciones críticas: resultados principales, cifras, referencias, fórmulas, decisiones metodológicas y recomendaciones con consecuencias prácticas. Cada elemento puede clasificarse como confirmado, parcialmente confirmado, contradictorio o no verificable. Después, debe conservarse un registro de la fuente consultada, la discrepancia detectada y la corrección realizada. Este procedimiento evita depender de una revisión general del texto, que suele pasar por alto errores pequeños pero decisivos. Además, facilita que otros investigadores reconstruyan cómo se controlaron las salidas generadas durante el estudio.

La respuesta generativa debe tratarse como un producto provisional y no como evidencia científica. La herramienta puede acelerar búsquedas, explicaciones, cálculos o redacción, pero cada salida necesita un nivel de control proporcional a su importancia. La verificación rigurosa combina fuentes primarias, revisión conceptual, reproducción de

operaciones, triangulación y documentación de cambios. Este protocolo no elimina por completo los errores, pero reduce la posibilidad de que información inventada o razonamientos incorrectos ingresen en el manuscrito y se propaguen hacia análisis, conclusiones o decisiones posteriores.

Figura 13

Protocolo de verificación de resultados generativos en seis pasos



Nota. La figura presenta un protocolo operativo para verificar información producida por sistemas generativos. La revisión debe priorizar afirmaciones críticas y utilizar evidencia externa independiente, pues la coherencia textual y la confianza expresada por el modelo no demuestran exactitud.

La verificación de resultados generativos debe desarrollarse como un procedimiento sistemático que permita distinguir entre información confirmada, corregida, contradictoria o no verificable. La revisión de fuentes primarias, la comprobación bibliográfica, la reproducción de cálculos y código, y la triangulación independiente reducen el riesgo de incorporar errores plausibles al trabajo científico. La calidad de una respuesta no depende de su fluidez ni de la seguridad con que se presenta, sino de la evidencia que la respalda y de la trazabilidad de las decisiones adoptadas.

4.6 Integridad científica, autoría y declaración del uso de inteligencia artificial

La integridad científica comprende los principios y prácticas que garantizan honestidad, rigor, transparencia y responsabilidad durante la producción y comunicación del conocimiento. La incorporación de inteligencia artificial amplía las posibilidades de apoyo, pero también introduce riesgos relacionados con atribución, originalidad, confidencialidad y control de resultados. Masbernat (2024) sostiene que estas herramientas deben utilizarse dentro de marcos que permitan verificar las decisiones adoptadas. La responsabilidad final permanece en las personas que diseñan el estudio, revisan las salidas y presentan el trabajo ante la comunidad científica.

La autoría científica no depende únicamente de producir fragmentos de texto, código o imágenes. Supone realizar contribuciones intelectuales identificables, aprobar la versión final y responder por la exactitud e integridad del contenido. Por esta razón, un sistema de inteligencia artificial no puede asumir plenamente la condición de autor, ya que no responde ante observaciones editoriales, conflictos éticos o errores descubiertos después de la publicación. Alonso (2024) plantea que el uso académico de sistemas generativos obliga a revisar los criterios tradicionales de atribución sin desplazar la responsabilidad que corresponde a los investigadores humanos.

La responsabilidad autoral incluye verificar datos, citas, procedimientos, análisis y conclusiones, aunque hayan sido producidos o modificados con asistencia automatizada. Tampoco puede atribuirse a la herramienta una decisión metodológica equivocada o una referencia inexistente. Los autores deben comprender suficientemente el contenido para defenderlo, corregirlo y explicar cómo fue obtenido. Tang et al. (2025) advierten que el aumento de las capacidades autónomas exige priorizar salvaguardas, control humano y rendición de cuentas. Delegar una tarea no implica transferir la obligación de responder por sus consecuencias científicas.

El plagio puede aparecer cuando un sistema reproduce expresiones, argumentos o estructuras procedentes de materiales existentes sin una atribución adecuada. Aunque el usuario no haya copiado directamente una fuente, sigue siendo responsable de comprobar la originalidad del texto entregado. Barrios (2023) analiza las tensiones que la inteligencia artificial genera en torno a la escritura y la producción académica. Los investigadores deben revisar coincidencias, reconocer ideas ajenas y citar las fuentes primarias

consultadas, evitando presentar una reformulación automatizada como una contribución intelectual propia y completamente original.

La fabricación consiste en inventar datos, observaciones, participantes o resultados que nunca existieron, mientras la falsificación implica modificar, omitir o manipular materiales y procedimientos para producir una representación engañosa. Los sistemas generativos pueden crear cifras, tablas o narraciones plausibles sin respaldo empírico. Cano et al. (2023) destacan la necesidad de utilizar estas tecnologías con criterios éticos y formativos. Ninguna salida sintética debe incorporarse como evidencia observada. Los datos, ejemplos simulados y resultados reales deben encontrarse claramente diferenciados y vinculados con registros verificables.

La confidencialidad requiere controlar qué información se introduce en plataformas externas. Entrevistas, historias clínicas, manuscritos inéditos, datos institucionales y documentos sometidos a revisión pueden contener elementos protegidos o permitir la identificación indirecta de personas. Comas (2025) señala que el uso de inteligencia artificial en contextos académicos plantea obligaciones de protección y manejo responsable de la información. Antes de cargar materiales, deben revisarse las políticas de almacenamiento, entrenamiento, conservación y transferencia. La anonimización debe realizarse previamente y no depender únicamente de controles ofrecidos por la herramienta.

La propiedad intelectual también debe considerarse cuando la inteligencia artificial interviene en la producción de textos, código, imágenes o materiales derivados. Los autores necesitan comprobar si los contenidos incorporados poseen licencias compatibles, si existen restricciones de reutilización y si la salida reproduce elementos protegidos. El apoyo automatizado no elimina las obligaciones de reconocimiento y autorización. Además, conviene distinguir entre ideas generales, expresiones originales, materiales de dominio público y contenidos con derechos vigentes. La revisión debe atender las políticas institucionales, contractuales y editoriales aplicables al proyecto científico.

El uso lingüístico comprende corrección gramatical, traducción, claridad y reducción de redundancias sin modificar el contenido sustantivo. Aunque suele representar una intervención limitada, debe revisarse para comprobar que no cambie conceptos, resultados o niveles de certeza. El uso sustantivo ocurre cuando la herramienta participa en formulación de hipótesis, generación de código, análisis, clasificación, interpretación

o redacción argumentativa. Messeri y Crockett (2024) advierten que estas tecnologías pueden producir una apariencia de comprensión que oculta limitaciones científicas. Por ello, el grado de intervención debe describirse con precisión.

La declaración del uso de inteligencia artificial debe identificar la herramienta, la versión disponible, la fecha de utilización y las tareas realizadas. También resulta conveniente explicar qué información se proporcionó, cómo fueron revisadas las salidas y qué decisiones permanecieron bajo control humano. Una frase genérica indicando que “se utilizó IA” no permite evaluar su influencia metodológica. Cuando la intervención afectó análisis, código o selección de evidencia, la declaración debe incorporarse en la metodología. Si se limitó a revisión lingüística, puede informarse en agradecimientos o en la sección exigida por la revista.

Las políticas editoriales pueden variar en cuanto al tipo de uso permitido, la ubicación de la declaración y la documentación requerida. Antes de someter un manuscrito, los autores deben revisar las instrucciones de la revista, las normas institucionales y los criterios del organismo financiador. Lu et al. (2026) muestran que los sistemas actuales pueden participar en generación de ideas, experimentación computacional, análisis y redacción, lo que vuelve insuficiente una declaración centrada exclusivamente en corrección textual. La transparencia debe corresponder con la profundidad real de la intervención tecnológica.

La rendición de cuentas exige conservar registros que permitan examinar cómo la herramienta influyó en el trabajo. Estos registros pueden incluir prompts, versiones, respuestas, modificaciones, verificaciones y responsables de cada decisión. No siempre será necesario publicar todas las interacciones, pero sí conservar evidencia suficiente para una auditoría o revisión posterior. Comas (2025) destaca que la transparencia fortalece la confianza cuando permite diferenciar contribuciones humanas y automatizadas. Esta documentación también facilita corregir el manuscrito si posteriormente se identifica una salida defectuosa o una interpretación indebidamente aceptada.

En consecuencia, la integridad científica con inteligencia artificial depende menos de prohibir herramientas que de establecer límites, controles y responsabilidades verificables. La asistencia puede ser válida cuando su finalidad está definida, sus resultados son revisados y su uso se declara conforme a las políticas aplicables. Los autores humanos conservan la obligación de garantizar originalidad, proteger

información, comprobar fuentes y responder por el contenido completo. La inteligencia artificial puede apoyar procesos científicos cada vez más amplios, pero no reemplaza la responsabilidad ética, intelectual y pública asociada con la autoría.

Tabla 16

Uso de inteligencia artificial, obligaciones de declaración y responsabilidad científica

Tipo de uso de IA	Obligación de declaración	Riesgo principal	Responsabilidad del autor
Corrección ortográfica y gramatical	Según la política editorial; recomendable cuando la intervención es amplia	Cambio involuntario del significado	Revisar que se conserven conceptos, cifras y niveles de certeza
Mejora de estilo y claridad	Informar cuando modifica de manera sustancial la redacción	Homogeneización de la voz o alteración argumentativa	Aprobar cada modificación y conservar precisión disciplinar
Traducción	Declarar la herramienta y la revisión realizada	Errores terminológicos o pérdida de matices	Validar la versión con competencia lingüística y disciplinar
Generación de títulos y palabras clave	Declaración breve cuando la política lo solicite	Representación imprecisa o exagerada del estudio	Seleccionar opciones coherentes con el contenido
Resumen de documentos	Declarar cuando influya en revisión, selección o interpretación	Omisión de limitaciones y atribución incorrecta	Comparar el resumen con la fuente primaria
Búsqueda de literatura	Registrar herramienta, fecha y alcance	Fuentes inexistentes, incompletas o desactualizadas	Verificar cada referencia y ampliar la búsqueda en bases confiables
Formulación de preguntas o hipótesis	Declaración metodológica sustantiva	Ideas plausibles sin fundamento o falta de originalidad	Evaluar pertinencia, novedad y comprobabilidad
Generación de instrumentos	Declarar el sistema y el proceso de validación	Ítems inválidos, ambiguos o culturalmente inadecuados	Validar contenido, comprensión y correspondencia con variables
Producción o modificación de código	Declarar herramienta, versión y nivel de intervención	Errores técnicos o transformaciones silenciosas	Inspeccionar, probar, documentar y reproducir el código
Análisis y clasificación de datos	Declaración detallada en metodología	Resultados incorrectos, sesgados o inestables	Validar métricas, casos, supuestos y resultados intermedios

Tipo de uso de IA	Obligación de declaración	Riesgo principal	Responsabilidad del autor
Elaboración de tablas, gráficos o imágenes	Declarar cuando el contenido o diseño sea generado sustantivamente	Distorsión visual, cifras erróneas o imágenes engañosas	Comprobar datos, escalas, etiquetas y correspondencia
Interpretación de resultados	Declaración metodológica obligatoria	Conclusiones no sustentadas o exceso de confianza	Contrastar con evidencia, teoría y revisión experta
Redacción sustantiva del manuscrito	Declarar herramienta, secciones afectadas y revisión humana	Pérdida de originalidad, alucinaciones o atribución indebida	Reescribir, comprobar y responder por el contenido completo
Revisión de manuscritos confidenciales	Solo cuando esté permitido; informar conforme a la política	Divulgación de contenido inédito o datos sensibles	No cargar materiales sin autorización y protección adecuada
Automatización de varias fases	Declaración integral del flujo, modelos, controles y decisiones	Propagación acumulativa de errores	Supervisar cada etapa, conservar registros y detener procesos inseguros

Nota. La tabla distingue usos lingüísticos y sustantivos de la inteligencia artificial en el trabajo científico. La obligación concreta de declaración depende de la política editorial e institucional, pero la revisión, la transparencia y la responsabilidad sobre el contenido final corresponden siempre a los autores humanos.

El uso responsable de inteligencia artificial en el trabajo científico exige distinguir entre apoyos lingüísticos y participaciones sustantivas que influyen en el diseño, el análisis o la interpretación. Cuanto mayor sea la intervención de la herramienta, más detallada debe ser su declaración y más rigurosos los controles aplicados. La transparencia no exime de responsabilidad: los autores deben verificar fuentes, datos, código, figuras y conclusiones, proteger la confidencialidad y conservar registros suficientes para demostrar cómo intervino la inteligencia artificial en el resultado final.

4.7 Revisión por pares y evaluación científica asistidas por inteligencia artificial

La revisión por pares asistida por inteligencia artificial comprende el empleo de sistemas automatizados para apoyar la evaluación formal, metodológica y documental de manuscritos científicos. Estas herramientas pueden ordenar verificaciones, señalar inconsistencias y preparar informes preliminares, reduciendo parte de la carga operativa de revisores y editores. Cárdenas (2023) reconoce que la inteligencia artificial está transformando las prácticas de comunicación y evaluación académica. Sin embargo, la aceptación, modificación o rechazo de una contribución debe conservarse como decisión humana sustentada en criterios científicos, editoriales y éticos explícitos.

En la evaluación formal, la inteligencia artificial puede comprobar extensión, estructura, formato de citas, presencia de secciones, correspondencia entre tablas y texto, y cumplimiento preliminar de instrucciones editoriales. Esta revisión permite detectar omisiones antes de iniciar la valoración especializada. Gimeno y Trigo (2024) señalan que las tecnologías automatizadas pueden contribuir a optimizar procesos vinculados con la publicación científica. Su aplicación resulta especialmente útil como filtro técnico, pero no permite determinar por sí sola si un manuscrito realiza una aportación original, relevante y suficientemente fundamentada para una disciplina.

La revisión metodológica puede apoyarse mediante listas de comprobación adaptadas al diseño del estudio. Un sistema puede identificar ausencia de criterios de inclusión, inconsistencias entre muestra y análisis, omisión de supuestos estadísticos o explicaciones insuficientes sobre instrumentos. También puede comparar métodos y resultados para localizar procedimientos no descritos. No obstante, una coincidencia con requisitos formales no demuestra validez metodológica. El revisor necesita examinar la pertinencia del diseño, la coherencia entre pregunta y evidencia, y las consecuencias de las decisiones adoptadas dentro del contexto disciplinar correspondiente.

La comprobación automatizada de referencias permite identificar citas ausentes en la bibliografía, registros duplicados, DOI incorrectos y posibles fuentes inexistentes. También puede señalar afirmaciones que requieren respaldo o referencias cuya fecha no corresponde con el texto. Montomoli (2023) destaca que la inteligencia artificial puede agilizar tareas editoriales y mejorar controles preliminares. Sin embargo, la herramienta no debe determinar la pertinencia intelectual de una fuente basándose únicamente en coincidencias textuales. El revisor debe consultar los documentos originales y comprobar si realmente respaldan la afirmación atribuida.

Los sistemas también pueden detectar contradicciones entre objetivos, métodos, resultados y conclusiones. Por ejemplo, pueden señalar variables anunciadas que no aparecen en el análisis, cifras diferentes entre tablas o interpretaciones que exceden los resultados. A partir de estas observaciones, es posible elaborar un informe preliminar organizado por secciones. Lu et al. (2026) muestran que los sistemas actuales pueden generar manuscritos y realizar evaluaciones automatizadas dentro de flujos científicos amplios. Este avance demuestra capacidad operativa, pero no convierte la evaluación automática en equivalente del juicio experto.

La confidencialidad constituye una limitación decisiva. Los manuscritos sometidos a evaluación contienen resultados inéditos, datos potencialmente sensibles e ideas protegidas temporalmente por el proceso editorial. Introducirlos en plataformas externas sin autorización puede vulnerar compromisos del revisor y políticas de la revista. Antes de utilizar una herramienta, deben comprobarse sus condiciones de almacenamiento, reutilización y acceso. El contenido solo debería procesarse mediante sistemas autorizados por la revista, con configuraciones adecuadas de privacidad y sin transferir información identificable o reservada a servicios no aprobados.

Los sesgos pueden influir cuando el sistema evalúa novedad, claridad, calidad lingüística o relevancia mediante patrones aprendidos de publicaciones predominantes. Esto puede favorecer enfoques, idiomas, regiones y estilos metodológicos con mayor presencia en sus datos. Un manuscrito innovador podría recibir una valoración desfavorable por apartarse de estructuras frecuentes. También pueden penalizarse trabajos interdisciplinarios o conocimientos locales. La revisión humana debe distinguir entre una deficiencia real y una diferencia respecto del patrón convencional, evitando que la automatización homogeneice los criterios de producción y comunicación científica.

La responsabilidad del revisor no desaparece cuando utiliza inteligencia artificial. Debe verificar cada observación, eliminar comentarios incorrectos y formular una evaluación respetuosa, específica y útil para los autores. Tang et al. (2025) advierten que los sistemas científicos autónomos pueden generar resultados aparentemente correctos que presentan fallas durante su aplicación, por lo que recomiendan priorizar salvaguardas sobre autonomía. En revisión por pares, esto implica impedir que un informe automático determine decisiones editoriales sin lectura experta, comprobación documental y análisis de las posibles consecuencias.

El modelo más adecuado es híbrido: la inteligencia artificial realiza comprobaciones repetitivas y organiza señales preliminares, mientras revisores y editores interpretan la contribución, ponderan discrepancias y deciden. Este enfoque puede disminuir carga operativa sin convertir la evaluación en un procedimiento mecánico. Deben registrarse las tareas automatizadas, las verificaciones humanas y las políticas aplicadas. La revisión por pares mantiene su legitimidad cuando existe confidencialidad, independencia y responsabilidad identificable. La eficiencia tecnológica resulta secundaria frente a la obligación de emitir decisiones fundamentadas, imparciales y científicamente defendibles.

Tabla 17*Tareas de la revisión por pares asistida por inteligencia artificial*

Tarea editorial	Posible apoyo de la inteligencia artificial	Riesgo principal	Decisión humana indelegable
Comprobación del formato	Revisar estructura, extensión, secciones y estilo bibliográfico	Rechazar contenidos valiosos por incumplimientos menores	Determinar si el manuscrito puede continuar a evaluación
Clasificación temática	Identificar área, palabras clave y posibles revisores	Asignación incorrecta por ambigüedad o interdisciplinariedad	Seleccionar editor y revisores con experiencia pertinente
Revisión metodológica inicial	Aplicar listas de comprobación y localizar omisiones	Confundir cumplimiento formal con validez científica	Evaluar pertinencia, rigor y coherencia del diseño
Comprobación de referencias	Detectar DOI incorrectos, citas faltantes y duplicados	Aceptar fuentes existentes que no respaldan la afirmación	Verificar contenido, pertinencia y calidad de las fuentes
Detección de inconsistencias	Comparar cifras, variables, tablas y conclusiones	Señalar como error una diferencia metodológicamente justificada	Determinar gravedad y efecto de la inconsistencia
Revisión estadística preliminar	Identificar supuestos omitidos o métricas incompatibles	Interpretar resultados sin conocer el diseño completo	Valorar adecuación del análisis y necesidad de correcciones
Elaboración de informe preliminar	Organizar observaciones por importancia y sección	Generar comentarios genéricos, falsos o contradictorios	Redactar y aprobar el informe enviado a los autores
Evaluación de originalidad	Comparar similitudes textuales y temáticas	Confundir terminología común con plagio o minimizar apropiaciones conceptuales	Examinar contexto, atribución e intención
Revisión lingüística	Detectar problemas de claridad, sintaxis y terminología	Penalizar variedades lingüísticas o estilos disciplinares	Decidir si la redacción impide comprender el estudio
Evaluación de relevancia	Comparar el manuscrito con	Favorecer temas dominantes y excluir enfoques novedosos	Valorar contribución, novedad y

Tarea editorial	Posible apoyo de la inteligencia artificial	Riesgo principal	Decisión humana indelegable
	literatura y alcance editorial		adecuación a la revista
Protección de confidencialidad	Procesar documentos en entornos editoriales autorizados	Filtración, almacenamiento o reutilización del manuscrito	Autorizar herramientas y proteger el contenido inédito
Recomendación editorial	Resumir fortalezas, debilidades y observaciones	Automatizar aceptación o rechazo mediante criterios opacos	Aceptar, solicitar cambios, rechazar o pedir otra evaluación

Nota. La tabla delimita tareas en las que la inteligencia artificial puede brindar apoyo durante la revisión editorial. La interpretación del aporte, la valoración metodológica, la protección del manuscrito y la recomendación final permanecen bajo responsabilidad de revisores y editores humanos.

La inteligencia artificial puede reducir parte de la carga operativa de la revisión por pares mediante comprobaciones formales, detección de inconsistencias, verificación preliminar de referencias y organización de observaciones. Sin embargo, estas funciones no sustituyen el juicio editorial ni la evaluación experta de la relevancia, originalidad y solidez metodológica del manuscrito. La confidencialidad, la interpretación contextual y la recomendación final deben permanecer bajo responsabilidad humana. El modelo más adecuado es híbrido: automatización para tareas repetitivas y decisión humana para las valoraciones científicas sustantivas.

Capítulo

5

Nuevos paradigmas de investigación en la era digital

Alexander Patricio Mosquera Urbano



5.1 Ciencia abierta y ecosistemas de datos reutilizables

La ciencia abierta busca que los conocimientos, materiales y procesos de investigación puedan consultarse, evaluarse y reutilizarse bajo condiciones transparentes. Su alcance supera la publicación gratuita de artículos, pues incluye datos, código, protocolos, instrumentos, resultados negativos y documentación metodológica. Choque et al. (2025) relacionan este enfoque con la democratización del conocimiento y el fortalecimiento de la colaboración científica. Cuando estos componentes se articulan adecuadamente, otros investigadores pueden comprobar resultados, desarrollar análisis adicionales y formular nuevas preguntas sin comenzar cada proyecto desde cero.

Las publicaciones abiertas constituyen una puerta de acceso al ecosistema científico, pero su disponibilidad no garantiza por sí sola la reutilización. Un artículo puede describir conclusiones sin proporcionar los datos, procedimientos o programas necesarios para reconstruirlas. Rodríguez et al. (2025) destacan que la apertura debe acompañarse de prácticas institucionales capaces de facilitar circulación y aprovechamiento del conocimiento. Por ello, las publicaciones necesitan conectarse con materiales complementarios mediante enlaces persistentes, identificadores y descripciones suficientemente claras para comprender qué recursos fueron utilizados y bajo qué condiciones pueden consultarse.

Los datos reutilizables deben conservar información sobre su procedencia, estructura, significado y proceso de creación. Compartir archivos sin variables definidas, unidades, categorías o criterios de exclusión limita su utilidad científica. También es necesario diferenciar datos originales, bases procesadas y conjuntos utilizados en cada análisis. Van et al. (2024) proponen una organización reproducible que vincula datos, código, resultados y entornos computacionales. Esta conexión permite que la reutilización no dependa de interpretaciones informales, sino de una documentación capaz de reconstruir las transformaciones realizadas durante el estudio.

El código y los protocolos representan componentes esenciales porque muestran cómo fueron procesados los datos y ejecutados los procedimientos. Los repositorios pueden almacenar scripts, cuadernos computacionales, configuraciones, instrucciones experimentales y registros de versiones. Sin embargo, publicar código sin dependencias, comentarios o ejemplos de ejecución dificulta su aplicación. Los protocolos deben indicar materiales, secuencias, decisiones y condiciones que afectan el resultado. La apertura

resulta científicamente valiosa cuando permite comprender el procedimiento, identificar sus límites y adaptarlo responsablemente a otros contextos investigativos.

Los metadatos describen los recursos disponibles y facilitan su localización e interpretación. Pueden incluir autoría, fecha, disciplina, licencia, formato, variables, instrumentos, población y restricciones de acceso. Sin esta información, un conjunto de datos puede permanecer técnicamente disponible, pero resultar prácticamente invisible o incomprensible. La calidad de los metadatos debe mantenerse durante todo el ciclo de vida del proyecto. Su estandarización favorece que buscadores, repositorios y sistemas de inteligencia artificial identifiquen relaciones entre publicaciones, datos, código y protocolos sin confundir recursos similares o versiones diferentes.

La interoperabilidad permite que datos y herramientas procedentes de sistemas distintos puedan relacionarse y utilizarse conjuntamente. Requiere formatos compatibles, vocabularios controlados, identificadores persistentes y reglas compartidas para describir variables y entidades. Gamella et al. (2025) muestran el potencial de integrar datos científicos heterogéneos para descubrir relaciones y apoyar inferencias más amplias. No obstante, la compatibilidad técnica no debe alcanzarse mediante la eliminación de diferencias relevantes. La integración necesita conservar el contexto disciplinar y documentar las equivalencias utilizadas para evitar interpretaciones incorrectas entre fuentes.

La reutilización implica que los materiales puedan verificarse, combinarse o aplicarse a nuevas preguntas sin perder su significado original. Para ello, deben existir licencias claras, documentación suficiente y condiciones de acceso comprensibles. Un recurso abierto no es automáticamente reutilizable si posee formatos obsoletos, dependencias inaccesibles o restricciones ambiguas. Shao et al. (2025) destacan que los sistemas colaborativos humano-IA pueden organizar flujos científicos complejos y facilitar la reutilización de procedimientos. Esta capacidad aumenta cuando los recursos están estructurados y conectados mediante estándares consistentes.

Los ecosistemas abiertos también favorecen colaboración entre instituciones, disciplinas y regiones. Investigadores diferentes pueden aportar datos, revisar código, mejorar protocolos o reproducir análisis, creando procesos acumulativos de conocimiento. La inteligencia artificial puede ayudar a localizar materiales compatibles, resumir documentación y proponer conexiones entre repositorios. Sin embargo, la colaboración

necesita reglas de reconocimiento, atribución y responsabilidad. Quienes producen datos y herramientas deben recibir crédito por su trabajo, mientras quienes reutilizan materiales deben informar modificaciones, limitaciones y nuevas condiciones de aplicación.

La apertura debe equilibrarse con privacidad, confidencialidad y consentimiento. No todos los datos pueden publicarse íntegramente, especialmente cuando contienen información personal, clínica, comunitaria o institucional sensible. En estos casos pueden utilizarse anonimización, acceso controlado, entornos seguros o versiones agregadas. El consentimiento debe especificar si la información podrá reutilizarse y bajo qué finalidades.

Las brechas de infraestructura, conectividad, idioma y financiamiento condicionan quién puede producir y reutilizar conocimiento abierto. Los repositorios, estándares y recursos computacionales suelen concentrarse en instituciones con mayor capacidad técnica, lo que puede ampliar desigualdades existentes. Choque et al. (2025) y Rodríguez et al. (2025) permiten comprender la apertura como una práctica vinculada con inclusión y cooperación, no solo con disponibilidad digital. Un ecosistema verdaderamente reutilizable requiere formación, soporte institucional, reconocimiento académico y mecanismos que faciliten la participación de comunidades científicas diversas.

Figura 14

Ecosistema de ciencia abierta y reutilización de resultados científicos



Nota. La figura representa la ciencia abierta como un ecosistema conectado de publicaciones, datos, código, protocolos, infraestructura y comunidades investigadoras. La reutilización requiere documentación, interoperabilidad, condiciones éticas y mecanismos que permitan reconstruir y adaptar los recursos disponibles.

La ciencia abierta adquiere mayor valor cuando publicaciones, datos, código, protocolos e infraestructuras se conectan dentro de un ecosistema reutilizable. Esta articulación facilita la verificación, la reproducción y el desarrollo de nuevas investigaciones, pero

exige metadatos consistentes, identificadores persistentes y condiciones claras de acceso. La reutilización científica también depende de licencias, atribución, privacidad, consentimiento y sostenibilidad institucional. Por ello, abrir recursos no consiste únicamente en publicarlos, sino en documentarlos y organizarlos para que puedan comprenderse, evaluarse y adaptarse responsablemente.

5.2 Investigación basada en agentes y sistemas multiagente

La investigación basada en agentes organiza tareas científicas mediante sistemas capaces de percibir información, planificar acciones, utilizar herramientas y comunicar resultados. En una arquitectura multiagente, estas capacidades se distribuyen entre componentes especializados que cooperan para resolver un problema común. El propósito no consiste en reproducir todo el ciclo científico dentro de un único modelo, sino en asignar funciones diferenciadas y coordinar sus aportes. Boiko et al. (2023) mostraron que la conexión entre modelos de lenguaje, búsqueda, programación y herramientas experimentales permite desarrollar sistemas científicos compuestos con capacidades operativas amplias.

La especialización permite que cada agente opere con instrucciones, recursos y criterios adaptados a una tarea determinada. Un agente puede localizar literatura, otro formular hipótesis, otro revisar datos y un cuarto evaluar la coherencia de las propuestas. Esta distribución reduce la exigencia de que un solo sistema domine simultáneamente todas las actividades. Ghareeb et al. (2026) presentan Robin, una arquitectura multiagente que articula generación de hipótesis y análisis de datos en biología experimental. El valor de este diseño proviene de combinar capacidades complementarias, no simplemente de aumentar el número de modelos participantes.

Los roles deben definirse mediante objetivos, entradas, salidas y límites de actuación. Un agente de literatura puede recuperar antecedentes y organizar evidencia; uno de hipótesis, producir explicaciones comprobables; y uno de análisis, ejecutar procedimientos sobre conjuntos de datos. La claridad funcional evita duplicaciones y permite localizar el origen de cada resultado. Cornelio et al. (2023) muestran que los sistemas computacionales pueden combinar datos con conocimiento teórico para descubrir relaciones científicas. En una estructura distribuida, esta integración requiere especificar qué agente aporta evidencia, cuál interpreta relaciones y cuál comprueba su consistencia.

La coordinación establece cómo se asignan tareas, se administran prioridades y se resuelven dependencias. Puede existir un agente supervisor que descomponga el problema, distribuya solicitudes y determine cuándo una actividad debe repetirse. En Co-Scientist, un supervisor administra una cola de trabajo, asigna agentes especializados y distribuye recursos dentro de un marco asincrónico. Esta organización permite que varias funciones operen de forma continua sin seguir necesariamente una secuencia rígida. Sin embargo, el coordinador necesita criterios explícitos para evitar trabajos redundantes, interrupciones prematuras o concentración excesiva de decisiones en un solo componente.

La comunicación entre agentes debe transmitir resultados junto con su procedencia, nivel de confianza y condiciones de producción. Si un agente comparte únicamente una conclusión resumida, los demás pueden aceptarla sin conocer las fuentes, transformaciones o incertidumbres involucradas. Por ello, los mensajes internos deberían incorporar referencias, datos utilizados, parámetros y justificaciones. La interoperabilidad también resulta indispensable, porque cada componente puede manejar formatos diferentes. Una comunicación eficaz no se limita al intercambio de lenguaje natural; requiere estructuras que permitan verificar, comparar y reutilizar los productos generados por cada función especializada.

El debate multiagente busca mejorar las propuestas mediante generación, crítica y reformulación sucesivas. En lugar de aceptar la primera hipótesis, varios agentes pueden señalar debilidades, formular alternativas y evaluar cuál posee mayor respaldo. Gottweis et al. (2026) describen un sistema donde agentes especializados generan, critican, clasifican y refinan hipótesis mediante un proceso evolutivo de comparación. Esta dinámica puede ampliar la exploración intelectual, aunque el desacuerdo automatizado no garantiza profundidad. Los argumentos deben contrastarse con evidencia externa para impedir que la persuasión lingüística sea confundida con solidez científica.

La memoria compartida conserva documentos, decisiones, resultados parciales y observaciones producidas por los agentes. Su función consiste en mantener continuidad y evitar que cada componente repita búsquedas o ignore correcciones anteriores. Puede organizarse mediante bases estructuradas, repositorios documentales o registros de interacción. Sin embargo, una memoria común también facilita la propagación de errores. Si una afirmación incorrecta se almacena como conocimiento confirmado, puede

condicionar hipótesis, análisis y evaluaciones posteriores. Por esta razón, cada entrada debe incluir fuente, versión, fecha, responsable y estado de verificación.

El acceso a herramientas amplía las capacidades de los agentes más allá de la generación textual. Estos pueden consultar bases bibliográficas, ejecutar código, analizar archivos, utilizar modelos especializados o interactuar con instrumentos autorizados. Boiko et al. (2023) demostraron que un sistema compuesto puede combinar búsqueda documental, ejecución de código y automatización experimental para planificar y realizar procedimientos químicos. Las herramientas deben gestionarse mediante permisos, registros y límites operativos. Un agente no debería modificar datos, ejecutar acciones costosas o acceder a información sensible sin controles proporcionales al riesgo de la tarea.

La evaluación mutua permite que un agente revise la salida de otro antes de incorporarla a una respuesta conjunta. Pueden compararse referencias, cálculos, hipótesis, código y correspondencia entre evidencia y conclusión. Lu et al. (2026) integran generación, experimentación, análisis, redacción y revisión dentro de un sistema automatizado de investigación en aprendizaje automático. Sin embargo, la evaluación interna presenta límites cuando los componentes comparten modelos, datos o patrones de razonamiento. La coincidencia entre agentes relacionados no constituye validación independiente y debe complementarse con pruebas externas y revisión experta.

El consenso puede alcanzarse mediante votación, puntuación, jerarquización o decisión de un agente coordinador. Estas estrategias ayudan a seleccionar alternativas, pero una mayoría automatizada puede coincidir por compartir las mismas fuentes o limitaciones. La calidad del consenso depende de la diversidad real de enfoques, de la independencia de las evaluaciones y de la posibilidad de conservar posiciones minoritarias. En sistemas científicos, no siempre conviene eliminar el desacuerdo, porque una hipótesis menos favorecida puede representar una alternativa novedosa. El resultado debe mostrar criterios de selección, objeciones pendientes y grado de incertidumbre.

La participación humana debe mantenerse como componente estructural del sistema multiagente. Los investigadores definen objetivos, asignan funciones, autorizan herramientas y evalúan si las propuestas poseen relevancia disciplinar. También pueden intervenir cuando los agentes discrepan, cuando la evidencia es insuficiente o cuando una acción implica consecuencias éticas y materiales. Los sistemas recientes demuestran

capacidades relevantes de coordinación científica, pero continúan presentándose como colaboradores o infraestructuras experimentales, no como sustitutos consolidados de comunidades investigadoras. Su valor dependerá de combinar distribución funcional, trazabilidad, evaluación independiente y supervisión humana efectiva.

Figura 15

Organización distribuida de un sistema multiagente para la investigación científica



Nota. La figura representa una arquitectura multiagente centrada en la distribución de funciones científicas. La coordinación permite combinar agentes especializados, pero la memoria, el debate y el consenso deben someterse a verificación para evitar errores compartidos.

La organización distribuida permite que agentes especializados colaboren en la búsqueda de literatura, generación de hipótesis, preparación de datos, análisis, crítica y síntesis de resultados. Esta división funcional puede aumentar la eficiencia y ampliar la exploración científica, pero también puede propagar errores cuando los agentes comparten fuentes, supuestos o información incorrecta. Por ello, la memoria compartida, la trazabilidad, la evaluación mutua y el control de herramientas deben acompañarse de verificación

independiente. La decisión humana continúa siendo indispensable para resolver desacuerdos, validar la evidencia y autorizar la continuidad del proceso.

5.3 Laboratorios autónomos y ciencia de ciclo cerrado

La ciencia de ciclo cerrado integra diseño experimental, ejecución, medición, interpretación y aprendizaje dentro de un proceso iterativo. A diferencia de la automatización lineal, el sistema no se limita a repetir una secuencia establecida, sino que utiliza los resultados obtenidos para decidir qué experimento realizar después. Abolhasani y Kumacheva (2023) describen los laboratorios autodirigidos como entornos que combinan aprendizaje automático, automatización experimental y toma de decisiones. Su propósito es recorrer espacios de investigación extensos mediante iteraciones informativas, manteniendo objetivos, restricciones y criterios científicos previamente definidos.

El ciclo comienza con la formulación de un objetivo medible. Este puede consistir en maximizar el rendimiento de una reacción, descubrir un material con propiedades determinadas, reducir costos o comprender la relación entre variables experimentales. El objetivo debe expresarse mediante indicadores que el sistema pueda evaluar, pero también debe conservar significado científico. Una función matemática mal definida puede orientar la búsqueda hacia resultados técnicamente favorables y científicamente irrelevantes. Por ello, los investigadores deben establecer qué se optimizará, qué restricciones serán obligatorias y qué resultados requerirán interpretación adicional antes de continuar.

La etapa de diseño convierte el objetivo en un conjunto de experimentos posibles. El sistema determina variables, rangos, combinaciones y condiciones que podrían aportar información. En lugar de explorar exhaustivamente todas las alternativas, puede emplear modelos predictivos para priorizar regiones prometedoras. Xie et al. (2023) explican que la experimentación autónoma en materiales depende de integrar decisiones computacionales con procesos de laboratorio. El diseño debe reconocer limitaciones físicas, costos, disponibilidad de materiales y compatibilidad entre procedimientos, evitando formular experimentos que no puedan ejecutarse o interpretarse bajo las condiciones reales del laboratorio.

El aprendizaje activo orienta la selección del experimento siguiente según la incertidumbre del modelo y el valor esperado de la información. Puede priorizar condiciones donde las predicciones son inciertas, donde existe mayor probabilidad de mejorar el resultado o donde se necesita distinguir entre explicaciones alternativas. Esta estrategia permite reducir el número de pruebas necesarias. No obstante, su desempeño depende de los datos iniciales y de la representación del espacio experimental. Si el modelo parte de información limitada, puede concentrarse prematuramente en regiones conocidas y excluir posibilidades científicamente valiosas.

La optimización busca mejorar progresivamente una o varias propiedades mediante la comparación entre resultados anteriores y nuevas predicciones. Puede aplicarse a rendimiento, pureza, estabilidad, composición o consumo de recursos. Sin embargo, la búsqueda de un valor máximo no siempre coincide con el descubrimiento científico. Un sistema podría identificar condiciones eficientes sin explicar por qué funcionan o sin revelar relaciones generalizables. Tom et al. (2024) señalan que los laboratorios autodirigidos abarcan diversas tecnologías y objetivos, desde optimización de procesos hasta descubrimiento de materiales, lo que exige criterios de evaluación adaptados a cada finalidad.

Después del diseño, la ejecución transforma las condiciones seleccionadas en procedimientos experimentales concretos. La información sobre cantidades, tiempos, secuencias y configuraciones debe transmitirse de manera inequívoca a la infraestructura automatizada. Boiko et al. (2023) mostraron que Coscientist puede planificar y coordinar tareas químicas utilizando modelos de lenguaje y herramientas especializadas. Este caso evidencia que las instrucciones pueden generarse dinámicamente, pero también que deben respetarse límites operativos. Una propuesta computacional no debe ejecutarse cuando contradice protocolos de seguridad, excede capacidades instrumentales o utiliza materiales no autorizados.

La medición produce la evidencia que permite evaluar cada experimento. Los resultados deben registrarse mediante procedimientos calibrados y asociados con las condiciones exactas de ejecución. No basta con obtener un valor final; también deben conservarse señales intermedias, incidencias y posibles desviaciones. En A-Lab, Szymanski et al. (2023) combinaron planificación computacional, síntesis automatizada y caracterización para examinar materiales inorgánicos. El estudio muestra la importancia de conectar la

producción experimental con métodos capaces de verificar qué producto se obtuvo y si corresponde con la predicción formulada.

La interpretación convierte las mediciones en información útil para actualizar el conocimiento del sistema. Puede incluir clasificación del resultado, estimación de propiedades, comparación con predicciones y detección de anomalías. Esta fase no debe confundirse con una simple lectura instrumental, porque requiere decidir qué significa el resultado respecto del objetivo. Una medición inesperada podría representar un error técnico, una muestra contaminada o un fenómeno novedoso. Antes de incorporarla al modelo, deben revisarse su calidad, consistencia y plausibilidad para impedir que una observación defectuosa modifique indebidamente la dirección de la investigación.

La actualización incorpora los nuevos datos al modelo y modifica sus estimaciones sobre el espacio experimental. El sistema puede ajustar relaciones entre variables, reducir incertidumbre y reorganizar la prioridad de las alternativas pendientes. Esta capacidad diferencia un ciclo cerrado adaptativo de una automatización que ejecuta listas fijas. Sin embargo, cada actualización también puede acumular errores de medición, etiquetado o interpretación. Por ello, conviene establecer controles sobre qué resultados ingresan al modelo, cómo se ponderan las observaciones atípicas y cuándo debe repetirse un experimento antes de aceptar su información como válida.

La selección del siguiente experimento reinicia el ciclo con conocimiento actualizado. El sistema puede elegir entre confirmar un resultado, explorar una región nueva o refinar una condición prometedora. La repetición continúa hasta alcanzar un objetivo, agotar recursos o cumplir un criterio de detención. Abolhasani y Kumacheva (2023) destacan que estos sistemas buscan acelerar el descubrimiento mediante decisiones adaptativas. No obstante, detenerse también constituye una decisión científica: una mejora pequeña puede no justificar nuevas pruebas, mientras una anomalía relevante puede requerir abandonar temporalmente la optimización y realizar experimentos explicativos.

La reproducibilidad exige conservar datos, modelos, condiciones, versiones y decisiones de cada iteración. Sin este registro, el laboratorio puede producir un resultado eficiente que otros equipos no puedan reconstruir. Tom et al. (2024) identifican la integración de hardware, software y gestión de datos como condición esencial para el desarrollo de laboratorios autodirigidos. Los protocolos deben permitir rastrear por qué se seleccionó cada experimento, cómo se ejecutó y qué información modificó el modelo. También

conviene repetir condiciones críticas para distinguir mejoras reales de variaciones instrumentales o resultados circunstanciales.

La seguridad y el control humano atraviesan todo el ciclo. Los investigadores deben definir sustancias prohibidas, límites de operación, reglas de interrupción y decisiones que requieren autorización expresa. También deben evaluar resultados inesperados y revisar los criterios de optimización. La ciencia de ciclo cerrado amplía la velocidad y continuidad de la experimentación, pero no convierte al sistema en responsable de sus acciones. Su aplicación resulta adecuada cuando las variables son medibles, los procedimientos pueden estandarizarse y existen mecanismos de supervisión. La autonomía útil depende de límites verificables, no de eliminar la participación humana.

Tabla 18

Beneficios, riesgos y condiciones de aplicación de la ciencia de ciclo cerrado

Aspecto	Beneficio potencial	Riesgo principal	Condición de aplicación
Diseño experimental	Explora numerosas combinaciones de forma estratégica	Objetivos formulados mal o variables relevantes omitidas	Función o científicamente justificada
Aprendizaje activo	Prioriza experimentos informativos	Concentración prematura en regiones conocidas	Equilibrio entre exploración y aprovechamiento
Optimización	Mejora propiedades con menos iteraciones	Optimizar indicadores sin valor científico suficiente	Criterios múltiples y revisión disciplinar
Ejecución automatizada	Aumenta velocidad y consistencia operativa	Aplicación de instrucciones inseguras o inviables	Límites físicos, materiales y protocolos autorizados
Medición	Genera retroalimentación inmediata	Incorporación de datos defectuosos	Calibración y controles de calidad
Interpretación	Compara resultados y predicciones rápidamente	Confundir anomalías con descubrimientos o errores	Revisión de plausibilidad y repetición de casos críticos
Actualización del modelo	Aprende de cada experimento realizado	Propagación acumulativa errores	Filtrado, versionado y auditoría de datos

Aspecto	Beneficio potencial	Riesgo principal	Condición de aplicación
Selección adaptativa	Decide la prueba siguiente según nueva evidencia	Repetición innecesaria o exclusión de alternativas	Criterios de detención y diversidad experimental
Reproducibilidad	Registra secuencias y condiciones experimentales	Dependencia de equipos configuraciones no documentadas	de Protocolos, código, o versiones y registros no completos
Autonomía experimental	Mantiene la investigación durante ciclos prolongados	Pérdida de supervisión sobre decisiones críticas	de Control humano, sobre alertas y mecanismos de interrupción

Nota. La tabla resume las ventajas y condiciones necesarias para aplicar ciclos experimentales autónomos. La eficiencia depende de objetivos verificables, mediciones confiables, documentación completa y supervisión humana proporcional al riesgo.

La ciencia de ciclo cerrado puede acelerar la experimentación mediante selección adaptativa, aprendizaje activo y actualización continua de los modelos. No obstante, su eficacia depende de objetivos científicamente justificados, mediciones confiables y criterios capaces de equilibrar exploración y optimización. La automatización también exige documentar cada iteración, controlar la calidad de los datos y establecer mecanismos de interrupción ante resultados anómalos o acciones inseguras. La autonomía experimental resulta pertinente cuando permanece acompañada de supervisión humana, trazabilidad y validación disciplinar.

5.4 Investigación multimodal y conocimiento integrado

La investigación multimodal combina fuentes heterogéneas para construir representaciones más amplias de un fenómeno. Puede integrar textos, imágenes, audios, videos, señales biológicas, datos tabulares y conocimiento disciplinar dentro de un mismo proceso analítico. Su propósito no consiste simplemente en acumular formatos, sino en relacionar evidencias que describen dimensiones diferentes del objeto estudiado. Truhn et al. (2024) muestran este potencial en oncología de precisión, donde la información clínica, molecular, textual y visual puede aportar perspectivas complementarias sobre una misma condición.

Cada modalidad registra aspectos particulares y posee limitaciones propias. El texto permite documentar antecedentes, diagnósticos o experiencias; las imágenes muestran estructuras espaciales; el audio conserva elementos de la expresión oral; el video incorpora secuencias e interacciones; y las señales representan cambios fisiológicos o

instrumentales. Los datos tabulares, por su parte, organizan variables previamente definidas. Girdwood y Manikonda (2026) destacan que el uso de inteligencia artificial debe conservar sensibilidad frente a las características del material estudiado, pues ninguna modalidad constituye por sí sola una representación completa del fenómeno.

La integración requiere establecer correspondencias entre fuentes que pueden utilizar unidades, escalas y estructuras diferentes. Una imagen clínica puede vincularse con un informe textual, resultados de laboratorio y antecedentes registrados en una tabla. Para realizar esta relación deben existir identificadores, referencias temporales y criterios que indiquen qué registros pertenecen al mismo caso. La coincidencia técnica no es suficiente: dos datos pueden estar asociados con una persona, pero corresponder a momentos o condiciones distintas. La integración debe conservar esas diferencias para impedir comparaciones metodológicamente inadecuadas.

Las representaciones compartidas permiten traducir distintas modalidades hacia espacios computacionales donde puedan relacionarse. En estos espacios, el sistema identifica asociaciones entre palabras, patrones visuales, señales y variables estructuradas. Truhn et al. (2024) explican que los modelos multimodales pueden combinar información textual y visual mediante arquitecturas relacionadas, ampliando sus posibles aplicaciones en investigación oncológica. Sin embargo, una representación común simplifica inevitablemente parte del contenido original. Por ello, los resultados integrados deben poder rastrearse hasta los registros concretos que participaron en su generación.

El conocimiento disciplinar orienta qué modalidades deben combinarse y cómo interpretar sus relaciones. Un sistema puede detectar una asociación estadística entre una imagen y una variable clínica, pero necesita conceptos especializados para determinar su significado. Ocampo et al. (2024) muestran que el análisis multimodal en contextos educativos requiere vincular datos de interacción con procesos pedagógicos y condiciones de aprendizaje. La misma lógica se aplica a otras áreas: la integración técnica adquiere valor científico cuando responde a preguntas conceptualmente justificadas y no únicamente a coincidencias computacionales.

La calidad desigual de las fuentes representa uno de los desafíos principales. Una modalidad puede contener mediciones precisas, mientras otra presenta registros incompletos, ruido o descripciones subjetivas. Si el sistema integra ambas sin diferenciar su confiabilidad, la fuente más abundante puede dominar el resultado aunque sea menos

pertinente. Truhn et al. (2024) señalan que los datos multimodales en oncología presentan diferencias de disponibilidad, estandarización y escala. La integración debe incorporar controles de calidad específicos para cada modalidad y documentar qué información estuvo ausente o fue descartada.

La alineación temporal también condiciona la interpretación. Los textos, imágenes, señales y variables pueden haberse recopilado en momentos diferentes, por lo que no siempre describen el mismo estado del fenómeno. En un estudio clínico, una imagen inicial no debería relacionarse automáticamente con una medición obtenida después de una intervención. En investigación educativa, un registro audiovisual puede corresponder a una actividad distinta de aquella descrita en un cuestionario. Las marcas temporales y las secuencias deben conservarse para evitar relaciones aparentes producidas por una integración descontextualizada.

Shao et al. (2025) presentan SciSciGPT como un colaborador capaz de organizar flujos analíticos complejos, ejecutar procedimientos diversos y facilitar iteraciones reproducibles. Estas capacidades pueden apoyar la exploración de conjuntos multimodales mediante generación de código, organización de archivos y conexión entre tareas. Sin embargo, el sistema no debe decidir por sí solo qué modalidades son equivalentes o qué relación posee significado científico. El investigador necesita revisar transformaciones, criterios de unión y resultados intermedios, especialmente cuando la integración depende de múltiples herramientas automatizadas.

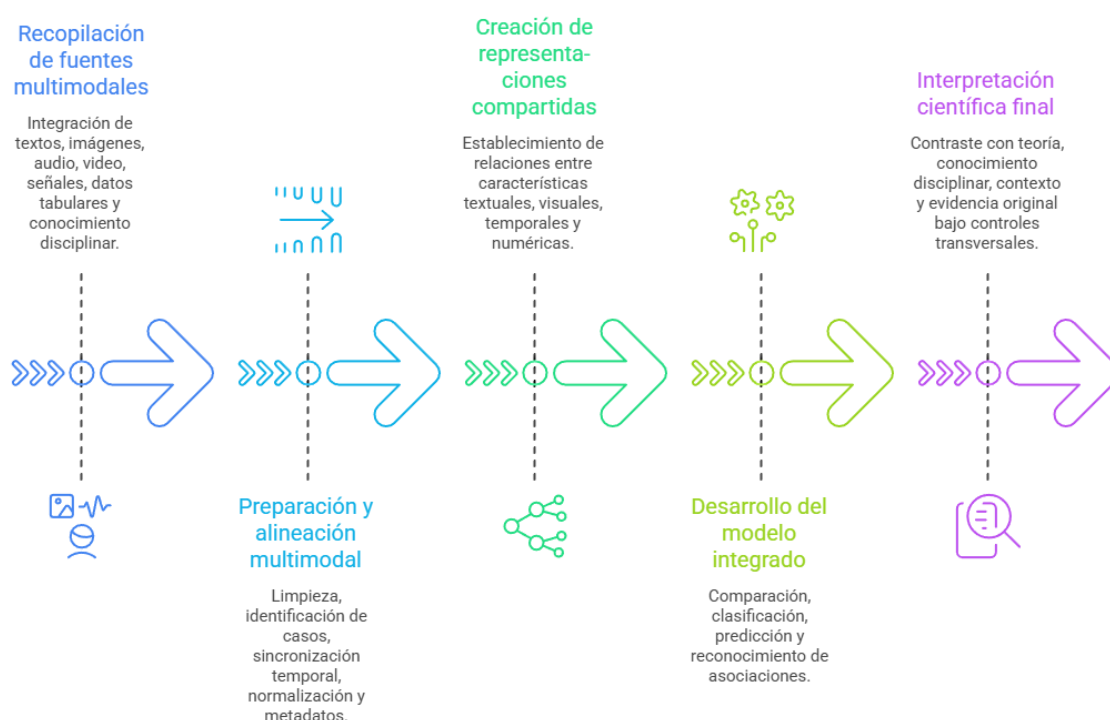
La trazabilidad de fuentes permite reconstruir qué registros sustentan una conclusión integrada. Cada resultado debe conservar vínculos con los documentos, imágenes, señales y variables utilizados, además de las transformaciones aplicadas. Esto resulta esencial cuando dos modalidades ofrecen resultados contradictorios. En lugar de ocultar la discrepancia mediante un promedio o una representación conjunta, debe examinarse si responde a diferencias temporales, errores o dimensiones distintas del fenómeno. La contradicción puede aportar conocimiento relevante y orientar nuevas preguntas, siempre que permanezca visible dentro del proceso analítico.

En consecuencia, la investigación multimodal produce conocimiento integrado cuando relaciona fuentes complementarias sin eliminar sus diferencias. Su calidad depende de representaciones compartidas, alineación temporal, documentación y participación disciplinar. La inteligencia artificial facilita el procesamiento conjunto de grandes

volúmenes de información, pero no reemplaza la interpretación contextual. El resultado final debe distinguir qué aporta cada modalidad, cómo se resolvieron las discrepancias y qué limitaciones permanecen. Integrar no significa fusionar indiscriminadamente, sino construir una interpretación verificable a partir de evidencias heterogéneas y científicamente pertinentes.

Figura 16

Convergencia de modalidades hacia un modelo integrado de conocimiento científico



Nota. La figura representa la integración de modalidades heterogéneas mediante procesos de alineación y construcción de representaciones compartidas. La interpretación científica requiere mantener la conexión con las fuentes originales, evaluar su calidad y aplicar conocimiento disciplinar.

La investigación multimodal permite integrar textos, imágenes, audios, videos, señales, datos tabulares y conocimiento disciplinar dentro de una representación científica conjunta. Sin embargo, esta convergencia solo resulta válida cuando cada modalidad conserva su procedencia, temporalidad y significado contextual. La preparación, alineación y comparación de fuentes deben acompañarse de controles de calidad y revisión de discrepancias.

5.5 Inferencia causal y descubrimiento de mecanismos

La inteligencia artificial puede identificar regularidades, formular hipótesis y orientar experimentos, pero estas capacidades no convierten automáticamente una asociación en

una relación causal. La inferencia causal busca establecer si modificar una variable produciría un cambio en otra bajo condiciones definidas. El descubrimiento de mecanismos intenta avanzar todavía más, explicando mediante qué procesos ocurre ese efecto. Krenn et al. (2022) sostienen que la contribución científica de la inteligencia artificial debe valorarse por su capacidad para favorecer comprensión, no únicamente por predecir resultados con exactitud.

La correlación indica que dos variables cambian conjuntamente, sin determinar por qué ocurre esa relación. Puede deberse a influencia directa, causalidad inversa, factores comunes o coincidencias producidas por la muestra. Los modelos de aprendizaje automático detectan asociaciones complejas y no lineales, pero su rendimiento predictivo no permite distinguir entre estas posibilidades. Una variable puede anticipar correctamente un resultado porque funciona como indicador indirecto de otra causa. Por ello, la correlación constituye una señal para investigar, no una demostración suficiente de que exista un vínculo causal entre los fenómenos observados.

La predicción responde qué resultado es probable dadas ciertas características disponibles. Su calidad se evalúa mediante desempeño en datos no utilizados durante el ajuste, estabilidad y capacidad de generalización. Un modelo puede predecir adecuadamente sin representar el proceso que genera el fenómeno. Incluso puede apoyarse en variables correlacionadas que perderían utilidad cuando cambian las condiciones. Krenn et al. (2022) distinguen entre resolver tareas mediante regularidades y contribuir a una comprensión transferible. La predicción adquiere valor explicativo solo cuando se conecta con hipótesis, mecanismos y pruebas independientes.

La causalidad responde qué ocurriría si una variable fuera modificada mediante una intervención. Esta pregunta difiere de observar qué sucede cuando la variable presenta valores distintos de manera natural. Las personas expuestas y no expuestas pueden diferir en múltiples características que también influyen sobre el resultado. La inferencia causal requiere definir la intervención, el resultado, la población y los factores que deben controlarse. Sin estas condiciones, el modelo puede atribuir efectos a una variable cuando la relación observada depende de elementos no medidos o seleccionados incorrectamente.

El razonamiento contrafactual compara un resultado observado con aquello que habría sucedido para la misma unidad bajo una condición alternativa. Como ambos escenarios no pueden observarse simultáneamente, el efecto causal debe estimarse mediante diseños

y supuestos. Los experimentos aleatorizados aproximan esta comparación al distribuir condiciones entre grupos comparables. En estudios observacionales se utilizan emparejamiento, ponderación, variables instrumentales u otros procedimientos, pero sus conclusiones dependen de supuestos que deben declararse. La inteligencia artificial puede apoyar estas estimaciones, aunque no elimina el problema fundamental de la evidencia contrafactual ausente.

Los grafos causales representan variables mediante nodos y posibles relaciones causales mediante flechas. Su utilidad consiste en explicitar supuestos, distinguir causas, mediadores y factores de confusión, y determinar qué variables deben ajustarse. Sin embargo, un grafo no constituye una descripción verdadera solo por estar formalizado. Puede construirse desde conocimiento disciplinar, datos, algoritmos o una combinación de fuentes. Cornelio et al. (2023) muestran que integrar principios teóricos con procedimientos computacionales puede orientar el descubrimiento de relaciones científicas, siempre que las estructuras propuestas se sometan a comprobación empírica.

Los datos observacionales permiten estudiar fenómenos cuando intervenir resulta imposible, costoso o éticamente inadecuado. Los algoritmos de descubrimiento causal pueden proponer estructuras compatibles con independencias estadísticas, cambios entre entornos o variaciones temporales. Sin embargo, diferentes grafos pueden explicar los mismos datos, especialmente cuando existen variables no observadas. Gamella et al. (2025) desarrollaron cámaras causales físicas para evaluar métodos mediante sistemas manipulables y mediciones controladas, destacando la importancia de probar algoritmos en escenarios donde las relaciones pueden contrastarse experimentalmente.

Los experimentos aportan evidencia más directa porque modifican deliberadamente una condición y observan sus consecuencias. No obstante, también presentan límites relacionados con cumplimiento, medición, tamaño muestral y generalización hacia otros contextos. La inteligencia artificial puede sugerir intervenciones informativas, priorizar experimentos y comparar resultados con predicciones causales. Las cámaras propuestas por Gamella et al. (2025) permiten manipular variables físicas y estudiar tareas como descubrimiento causal, generalización fuera de distribución y detección de cambios. Estos entornos conectan propuestas algorítmicas con pruebas controladas y repetibles.

El conocimiento previo resulta indispensable para restringir relaciones imposibles y orientar interpretaciones plausibles. Las leyes físicas, secuencias temporales, teorías

disciplinarios y resultados experimentales pueden impedir que un algoritmo acepte estructuras estadísticamente compatibles, pero científicamente absurdas. Cory et al. (2024) muestran que la incorporación de conocimiento y restricciones puede dirigir búsquedas computacionales hacia soluciones científicamente útiles. No obstante, el conocimiento previo también puede contener errores o supuestos heredados. Debe utilizarse como restricción revisable, no como criterio que impida considerar evidencia inesperada o mecanismos alternativos.

Un mecanismo describe la cadena de procesos mediante la cual una causa produce un efecto. Identificarlo exige más que estimar la magnitud de una relación; requiere localizar componentes, secuencias, condiciones y niveles de organización. Puede involucrar variables mediadoras, interacciones y procesos dinámicos que deben observarse o intervenir. La inteligencia artificial puede proponer mecanismos a partir de literatura y datos, pero estos permanecen como hipótesis. Gottweis et al. (2026) presentan Co-Scientist como un sistema que genera, debate y refina hipótesis utilizando herramientas y retroalimentación humana, no como una fuente autónoma de demostraciones causales.

La validación causal debe combinar análisis estadístico, experimentos, replicación y conocimiento especializado. Una estructura propuesta puede comprobarse mediante nuevas intervenciones, predicciones fuera de muestra o análisis en poblaciones diferentes. También deben evaluarse explicaciones alternativas y resultados negativos. La coincidencia entre varios agentes o modelos no representa confirmación independiente si comparten datos y supuestos. Los sistemas de generación de hipótesis pueden ampliar las alternativas examinadas, pero los científicos deben decidir cuáles poseen plausibilidad, relevancia y evidencia suficiente para justificar experimentos adicionales.

En consecuencia, el tránsito desde asociación hasta mecanismo requiere niveles crecientes de evidencia. La correlación describe covariación; la predicción anticipa resultados; la causalidad estima efectos de intervenciones; el contrafactual compara escenarios posibles; y el mecanismo explica cómo se produce el cambio. La inteligencia artificial puede apoyar cada nivel mediante detección de patrones, construcción de grafos, diseño experimental y generación de hipótesis. Sin embargo, ninguna salida predictiva demuestra causalidad por sí misma. La explicación científica surge cuando las propuestas resisten intervenciones, contrastes independientes y evaluación disciplinar.

Tabla 19*Diferencias entre asociación, predicción, causalidad y mecanismo científico*

Término	Pregunta que responde	Tipo de evidencia requerida	Limitación principal
Correlación	¿Qué variables cambian conjuntamente?	Datos observacionales y medidas de asociación	No distingue causalidad, causalidad inversa ni factores comunes
Predicción	¿Qué resultado es probable dadas ciertas entradas?	Desempeño en datos independientes y validación externa	Puede funcionar sin representar el proceso generador
Causalidad	¿Qué cambiaría al intervenir sobre una variable?	Experimentos o diseños observacionales con supuestos causales explícitos	Depende de la identificación adecuada de factores de confusión
Intervención	¿Qué ocurre al modificar deliberadamente una condición?	Manipulación controlada y comparación de resultados	Puede ser inviable, costosa o limitada a un contexto
Contrafactual	¿Qué habría ocurrido bajo una condición alternativa?	Diseño que permita estimar resultados no observados	El escenario alternativo nunca se observa para la misma unidad
Grafo causal	¿Qué relaciones causales se proponen entre las variables?	Supuestos disciplinares, datos y comprobaciones de independencia	Varios grafos pueden ser compatibles con los mismos datos
Mediación	¿A través de qué variable opera parte del efecto?	Secuencia temporal, modelo causal y medición de mediadores	Puede confundirse con factores comunes no observados
Mecanismo	¿Cómo y mediante qué procesos se produce el efecto?	Intervenciones, evidencia multiescala, teoría y replicación	Una narración plausible puede carecer de comprobación experimental
Explicación científica	¿Por qué ocurre el fenómeno y bajo qué condiciones?	Integración de causalidad, mecanismo, teoría y evidencia independiente	Puede ser provisional y limitada a determinados contextos
Comprensión	¿Puede el conocimiento utilizarse para anticipar, intervenir y transferir relaciones?	Explicaciones verificadas, generalización y control experimental	No se demuestra únicamente mediante rendimiento predictivo

Nota. La tabla diferencia niveles de análisis que suelen confundirse al utilizar inteligencia artificial. Los patrones y predicciones pueden orientar hipótesis causales, pero las afirmaciones sobre intervenciones y mecanismos requieren diseños, supuestos y validaciones adicionales.

La asociación, la predicción, la causalidad y el mecanismo científico corresponden a niveles distintos de conocimiento y no deben utilizarse como equivalentes. La correlación identifica variaciones conjuntas y la predicción anticipa resultados, pero ninguna demuestra por sí sola qué ocurriría ante una intervención. Las afirmaciones causales requieren supuestos explícitos, diseños adecuados y control de factores de confusión, mientras que los mecanismos exigen evidencia adicional sobre los procesos que producen el efecto. La inteligencia artificial puede orientar estas indagaciones, pero no reemplaza la validación experimental ni la explicación científica fundamentada.

5.6 Gobernanza responsable de la inteligencia artificial científica

La gobernanza responsable de la inteligencia artificial científica comprende las reglas, funciones y controles mediante los cuales instituciones e investigadores determinan cómo pueden utilizarse sistemas automatizados dentro del trabajo académico. No se limita a formular principios generales, sino que establece quién autoriza cada uso, qué riesgos deben evaluarse y cómo se responderá ante errores. Masbernat (2024) vincula la integridad científica con estructuras de responsabilidad y trazabilidad. Una gobernanza efectiva debe abarcar desde herramientas de apoyo lingüístico hasta sistemas capaces de seleccionar análisis, ejecutar experimentos o proponer conclusiones.

Los niveles de autonomía permiten ajustar los controles a la capacidad real del sistema. En un nivel asistido, la inteligencia artificial recomienda opciones que una persona revisa antes de aplicarlas. La semiautonomía permite ejecutar tareas delimitadas dentro de parámetros aprobados, mientras la autonomía elevada incorpora planificación, selección de herramientas y modificación de acciones según resultados. Tang et al. (2025) advierten que el aumento de autonomía científica debe acompañarse de salvaguardas proporcionales. Cada nivel requiere permisos, límites operativos, registros y criterios claros para solicitar intervención humana.

La asignación de responsabilidades debe evitar vacíos entre desarrolladores, investigadores, instituciones, proveedores y comités de supervisión. El investigador principal responde por la pertinencia metodológica y el uso de las salidas; el equipo técnico, por la configuración y mantenimiento; y la institución, por políticas, infraestructura y mecanismos de reclamación. Chaaban et al. (2024) destacan la necesidad

de preparar a los investigadores para reconocer limitaciones y obligaciones asociadas con la inteligencia artificial. Ningún actor debe asumir que otro verificará automáticamente los riesgos o corregirá las consecuencias de una decisión defectuosa.

La auditoría permite reconstruir cómo funcionó el sistema y qué decisiones influyeron en los resultados. Debe incluir versiones, datos utilizados, prompts, parámetros, permisos, registros de ejecución, intervenciones humanas y cambios realizados durante el proyecto. No basta con revisar el producto final, porque los errores pueden haberse introducido durante la preparación de datos, el análisis o la interpretación. Masbernat (2024) sostiene que la transparencia y la conservación de registros fortalecen la integridad científica. Las auditorías pueden ser internas, externas o activarse después de una incidencia relevante.

La evaluación de riesgos debe realizarse antes de utilizar la herramienta y actualizarse cuando cambien los datos, el modelo o el alcance del proyecto. Conviene valorar probabilidad, gravedad, reversibilidad y población afectada. Los riesgos pueden incluir conclusiones falsas, exposición de información, discriminación, daño experimental, pérdida de propiedad intelectual o dependencia excesiva del sistema. Messeri y Crockett (2024) advierten que la inteligencia artificial puede generar ilusiones de comprensión y promover monoculturas metodológicas. La evaluación debe considerar tanto fallas técnicas como efectos sobre la diversidad científica y la toma de decisiones.

La privacidad exige determinar qué información puede introducirse en plataformas automatizadas y bajo qué condiciones. Los datos personales, clínicos, comunitarios o institucionales requieren minimización, anonimización, acceso restringido y consentimiento compatible con los nuevos usos. Comas (2025) destaca que la utilización académica de inteligencia artificial debe respetar obligaciones de protección y confidencialidad. Las instituciones necesitan clasificar tipos de datos, aprobar entornos seguros y prohibir el ingreso de información sensible en servicios no autorizados. También deben definir plazos de conservación y procedimientos para responder ante filtraciones.

La equidad debe examinarse mediante evaluaciones desagregadas por grupos, idiomas, regiones y disciplinas. Un sistema puede mostrar buen rendimiento promedio y producir errores sistemáticos en poblaciones poco representadas. Por ello, los controles deben incluir pruebas de desempeño, revisión de etiquetas, análisis de cobertura y mecanismos para corregir exclusiones. Barrios (2023) señala que la incorporación de inteligencia

artificial transforma prácticas académicas y puede profundizar desigualdades existentes. La gobernanza debe evitar que la eficiencia tecnológica se obtenga a costa de invisibilizar perspectivas, comunidades o formas alternativas de producir conocimiento.

La seguridad adquiere especial importancia cuando la inteligencia artificial controla código, infraestructura o experimentos físicos. Antes de conceder acceso, deben establecerse listas de acciones permitidas, límites de recursos, validaciones previas y entornos aislados para tareas de alto riesgo. Lu et al. (2026) muestran que los sistemas recientes pueden participar en programación, experimentación y evaluación dentro de flujos científicos amplios. Esta capacidad exige separar funciones críticas, restringir credenciales y comprobar resultados antes de transferirlos hacia equipos, bases oficiales o procesos que produzcan consecuencias difíciles de revertir.

El control de experimentos debe incorporar puntos de autorización humana según el nivel de riesgo. Un sistema puede proponer condiciones, pero la ejecución debe bloquearse cuando intervengan sustancias peligrosas, parámetros fuera de rango o procedimientos no aprobados. También deben existir protocolos de interrupción capaces de detener acciones ante fallas, mediciones anómalas, pérdida de comunicación o comportamiento inesperado. Tang et al. (2025) recomiendan priorizar salvaguardas frente a una autonomía creciente. Los mecanismos de detención deben probarse periódicamente y no permanecer como simples disposiciones escritas.

Los actores institucionales necesitan coordinar políticas, formación, infraestructura y supervisión. Los comités de ética pueden valorar impacto sobre participantes; las unidades de tecnología, revisar seguridad y acceso; las bibliotecas, apoyar documentación y repositorios; y los editores, establecer reglas de declaración. Chaaban et al. (2024) subrayan que la competencia en inteligencia artificial incluye capacidades críticas y éticas, no solo manejo técnico. La gobernanza mejora cuando cada actor conoce su función, dispone de canales para comunicar incidentes y participa en revisiones periódicas de las herramientas autorizadas.

En consecuencia, gobernar la inteligencia artificial científica implica convertir principios en procedimientos verificables. Los niveles de autonomía deben vincularse con permisos, auditorías, evaluaciones de riesgo y puntos de control humano. La privacidad, equidad y seguridad necesitan indicadores, responsables y evidencias de cumplimiento. La institución debe conservar capacidad para suspender herramientas, investigar incidentes

y corregir prácticas. La responsabilidad final permanece en personas y organizaciones identificables, incluso cuando el sistema ejecute tareas complejas. Una gobernanza sólida no elimina la innovación, sino que establece condiciones para desarrollarla sin sacrificar rigor ni confianza científica.

Tabla 20

Actores y mecanismos de gobernanza de la inteligencia artificial científica

Actor	Responsabilidad principal	Mecanismo de control	Evidencia de cumplimiento
Investigador principal	Definir objetivos, aprobar usos y responder por resultados	Revisión metodológica y autorización de tareas críticas	Protocolo aprobado, registro de decisiones y firma de validación
Equipo de investigación	Verificar datos, salidas, código y conclusiones	Revisión cruzada y listas de comprobación	Bitácoras, resultados revisados y correcciones documentadas
Especialistas técnicos	Configurar modelos, permisos, entornos y seguridad	Control de acceso, pruebas y gestión de versiones	Archivos de configuración, registros de ejecución e informes técnicos
Institución	Establecer políticas, infraestructura y responsabilidades	Normativa interna, capacitación y supervisión periódica	Política vigente, registros de formación y evaluaciones institucionales
Comité de ética	Evaluar riesgos para participantes y comunidades	Revisión previa, seguimiento y condiciones de aprobación	Dictamen, consentimiento, medidas de protección e informes de seguimiento
Comité de seguridad	Controlar riesgos físicos, digitales y operativos	Análisis de amenazas y protocolos de interrupción	Pruebas de seguridad, incidentes registrados y planes de respuesta
Responsable de protección de datos	Garantizar privacidad, minimización y acceso legítimo	Clasificación de datos y evaluación de impacto	Matriz de datos, permisos, anonimización y registro de accesos
Proveedor del sistema	Informar capacidades, límites, versiones y tratamiento de datos	Documentación técnica y compromisos contractuales	Fichas del modelo, condiciones de servicio e historial de actualizaciones
Unidad de auditoría	Reconstruir decisiones y verificar cumplimiento	Auditorías internas o externas	Informes, hallazgos, acciones correctivas y seguimiento

Actor	Responsabilidad principal	Mecanismo de control	Evidencia de cumplimiento
Biblioteca o repositorio	Conservar datos, código, metadatos y versiones	Estándares de depósito y control documental	Identificadores persistentes, metadatos y registros de acceso
Revista o editorial	Definir usos permitidos y requisitos de declaración	Políticas editoriales y revisión del manuscrito	Declaración de IA, formularios autorales y decisiones documentadas
Organismo financiador	Condicionar recursos a prácticas responsables	Requisitos de gestión, evaluación y reporte	Plan de gestión de riesgos e informes de cumplimiento
Usuarios o comunidades afectadas	Comunicar impactos, errores y preocupaciones	Canales de reclamación y consulta	Registros de consultas, respuestas y medidas correctivas

Nota. La tabla distribuye responsabilidades entre los actores que intervienen en la gobernanza de la inteligencia artificial científica. El cumplimiento debe demostrarse mediante políticas, registros, auditorías, evaluaciones de riesgo y decisiones humanas documentadas.

La gobernanza de la inteligencia artificial científica requiere distribuir responsabilidades de manera clara entre investigadores, equipos técnicos, instituciones, comités, proveedores, editoriales y comunidades afectadas. Los principios generales solo adquieren valor cuando se traducen en mecanismos verificables, como controles de acceso, auditorías, evaluaciones de riesgo, registros de decisiones y protocolos de respuesta. La responsabilidad no puede diluirse entre actores ni trasladarse al sistema utilizado. Una gobernanza sólida exige evidencia de cumplimiento, supervisión continua y capacidad institucional para corregir, suspender o limitar usos que comprometan la integridad científica.

5.7 Perspectivas futuras de la investigación científica

Las perspectivas futuras de la investigación científica estarán determinadas por la relación entre capacidades automatizadas, organización institucional y juicio humano. Los modelos generativos ya apoyan búsqueda, análisis, programación y comunicación, mientras los sistemas agénticos conectan varias de estas tareas. Zhang et al. (2025) plantean que la integración de grandes modelos de lenguaje en el método científico debe mantenerse alineada con objetivos humanos y métricas claras de evaluación. El cambio esperado no será uniforme, porque cada disciplina posee datos, riesgos, infraestructuras y criterios de evidencia diferentes.

Los asistentes especializados representan el escenario de transformación más inmediato. Estos sistemas se concentran en actividades delimitadas, como organizar literatura, revisar código, preparar visualizaciones o comprobar referencias. Guerrero et al. (2025) consideran que la inteligencia artificial puede incrementar la eficiencia de diversas tareas investigativas cuando funciona como apoyo metodológico. En este escenario, el científico conserva la dirección del trabajo, selecciona las herramientas y valida cada resultado. Su viabilidad depende de competencias técnicas suficientes para formular instrucciones, detectar errores y decidir cuándo una salida automatizada no debe utilizarse.

La colaboración aumentada supone una relación más interactiva en la que el sistema propone alternativas, ejecuta procedimientos y recibe retroalimentación del investigador. Shao et al. (2025) presentan SciSciGPT como un prototipo abierto de colaboración humano-IA orientado a automatizar flujos, ampliar alternativas analíticas y favorecer la reproducibilidad. Esta modalidad puede reducir barreras técnicas y permitir que los equipos examinen más hipótesis o configuraciones. Sin embargo, requiere conservar registros, comparar resultados y evitar que la comodidad operativa transforme las recomendaciones del sistema en decisiones aceptadas sin evaluación crítica.

Los sistemas multiagente amplían la colaboración al distribuir funciones entre componentes dedicados a literatura, hipótesis, análisis, crítica y síntesis. Ghareeb et al. (2026) desarrollaron Robin para automatizar generación de hipótesis y análisis de datos en biología experimental, mientras otras arquitecturas recientes coordinan agentes mediante supervisores y memorias compartidas. Estos avances muestran que la especialización automatizada puede organizar tareas científicas complejas. No obstante, la coincidencia entre agentes no constituye validación independiente cuando comparten modelos, datos o supuestos, por lo que continúa siendo necesaria la contrastación externa.

La autonomía supervisada constituye un escenario intermedio entre colaboración y automatización integral. En este modelo, la inteligencia artificial puede planificar y ejecutar secuencias dentro de límites previamente autorizados, mientras las decisiones críticas requieren aprobación humana. Lu et al. (2026) presentan un sistema que conecta ideación, búsqueda bibliográfica, experimentación computacional, análisis, redacción y revisión en investigaciones de aprendizaje automático. Su alcance demuestra posibilidades relevantes, pero permanece concentrado en experimentos computacionales y no elimina la necesidad de comprobar calidad, novedad, seguridad y pertinencia científica.

La multimodalidad permitirá relacionar textos, imágenes, señales, videos, datos tabulares y conocimiento especializado dentro de modelos integrados. Esta convergencia puede favorecer explicaciones más amplias y decisiones basadas en evidencias complementarias. Vimos et al. (2024) destacan la influencia creciente de herramientas inteligentes en la transformación de los procesos investigativos. Sin embargo, los avances dependerán de resolver diferencias de calidad, temporalidad, escala y contexto entre modalidades. La integración futura deberá conservar la procedencia de cada fuente y permitir revisar cómo contribuyó al resultado obtenido.

Los laboratorios autónomos podrían acelerar descubrimientos en áreas donde los procedimientos sean medibles, repetibles y compatibles con automatización física. La selección adaptativa de experimentos permitirá recorrer espacios extensos y aprovechar continuamente las mediciones generadas. Sin embargo, la expansión de estos entornos dependerá de costos, seguridad, mantenimiento, interoperabilidad y capacidad para reproducir resultados en otras instalaciones. La autonomía experimental será más viable en tareas delimitadas que en problemas abiertos donde todavía deben reformularse conceptos, negociar objetivos o interpretar observaciones ambiguas mediante conocimiento contextual.

La ciencia abierta puede convertirse en la infraestructura que conecte asistentes, agentes y laboratorios. Publicaciones, datos, código, protocolos y registros interoperables permitirán que los sistemas localicen evidencias, repitan procedimientos y desarrollen nuevas investigaciones. Rodríguez et al. (2025) vinculan la apertura con colaboración y circulación del conocimiento. No obstante, el acceso desigual a repositorios, capacidad computacional y modelos avanzados puede concentrar los beneficios en pocas instituciones. La reutilización deberá acompañarse de reconocimiento, protección de datos, licencias claras y apoyo para comunidades con menores recursos tecnológicos.

Las competencias científicas también cambiarán. Los investigadores necesitarán comprender diseño metodológico, evaluación de modelos, gestión de datos, documentación, seguridad y comunicación de incertidumbre. Cárdenas (2023) advierte que la inteligencia artificial transforma prácticas académicas y exige nuevas capacidades críticas. No todos deberán convertirse en programadores expertos, pero sí reconocer qué decisiones pueden delegarse, qué controles deben aplicarse y cómo verificar resultados. La formación futura tendrá que combinar conocimientos disciplinares con alfabetización computacional, ética aplicada y experiencia en colaboración con sistemas automatizados.

La regulación deberá diferenciar niveles de riesgo y autonomía en lugar de aplicar reglas idénticas a todos los usos. Una corrección lingüística no requiere los mismos controles que un sistema capaz de modificar datos, ejecutar experimentos o recomendar decisiones clínicas. Tang et al. (2025) proponen priorizar salvaguardas frente al avance de los científicos artificiales, considerando entradas humanas, acciones del agente y consecuencias ambientales. Las instituciones necesitarán procedimientos de autorización, auditoría, interrupción, notificación de incidentes y revisión periódica, además de políticas actualizables ante cambios rápidos de modelos y proveedores.

Ningún escenario futuro puede considerarse inevitable. La asistencia especializada, la colaboración aumentada y la autonomía supervisada coexistirán según las condiciones de cada campo y proyecto. La automatización integral será técnicamente más factible en dominios computacionales o experimentales altamente estructurados, pero conservará límites asociados con comprensión, seguridad, acceso y responsabilidad. Las decisiones institucionales determinarán si estas tecnologías amplían la diversidad científica o fortalecen la concentración tecnológica. El futuro dependerá menos de sustituir investigadores que de diseñar relaciones verificables entre capacidades automáticas, conocimiento disciplinar y control humano.

Tabla 21

Escenarios futuros de incorporación de inteligencia artificial en la investigación científica

Escenario	Características principales	Oportunidades	Riesgos	Papel humano	Condiciones de viabilidad
Asistencia especializada	Herramientas separadas para búsqueda, redacción, código, análisis o visualización	Ahorro de tiempo, apoyo técnico y reducción de tareas repetitivas	Errores puntuales, dependencia y aceptación acrítica	Formula solicitudes, revisa salidas y decide su incorporación	Competencias básicas, verificación y políticas claras de uso
Colaboración aumentada	Interacción iterativa entre investigador y sistema durante varias tareas	Exploración de alternativas, mayor productividad y análisis ampliado	Homogeneización del razonamiento y pérdida de control metodológico	Dirige el proceso, aporta contexto y corrige propuestas	Trazabilidad, comparación de métodos y revisión experta
Sistemas multiagente	Distribución de funciones entre agentes especializados y coordinados	Paralelización, debate, evaluación mutua y manejo de flujos complejos	Consenso aparente, errores compartidos y coordinación opaca	Define roles, resuelve desacuerdos y valida resultados	Memoria auditable, diversidad de agentes y controles externos

Escenario	Características principales	Oportunidades	Riesgos	Papel humano	Condiciones de viabilidad
Autonomía supervisada	Ejecución de secuencias adaptativas dentro de límites autorizados	Mayor velocidad, continuidad operativa y optimización	Propagación de errores y acciones fuera del propósito científico	Autoriza etapas críticas, establece límites y puede interrumpir	Permisos graduados, registros, seguridad y criterios de detención
Automatización integral delimitada	Conexión de ideación, programación, experimentación, análisis y escritura	Producción rápida de estudios en dominios estructurados	Resultados superficiales, fallas acumulativas y evaluación no independiente	Comprueba relevancia, originalidad, validez y comunicación	Problemas bien definidos, datos adecuados y validación externa
Investigación multimodal	Integración de texto, imagen, señales, video y datos estructurados	Representaciones más completas y hallazgos interdisciplinarios	Pérdida de contexto, calidad desigual y relaciones espurias	Interpreta cada modalidad y justifica su integración	Metadatos, alineación, trazabilidad y conocimiento disciplinar
Laboratorios autónomos	Selección, ejecución y actualización adaptativa de experimentos	Exploración eficiente y reducción del tiempo de descubrimiento	Riesgos físicos, optimización inadecuada y dependencia instrumental	Define objetivos, seguridad y criterios de validación	Infraestructura compatible, protocolos y control de emergencia
Ecosistemas científicos abiertos	Conexión de publicaciones, datos, código, protocolos y agentes	Reutilización, replicación y colaboración internacional	Brechas de acceso, privacidad y concentración de infraestructura	Produce, documenta, reutiliza y reconoce contribuciones	Interoperabilidad, licencias, protección de datos y sostenibilidad
Gobernanza adaptativa	Regulación proporcional al riesgo y al nivel de autonomía	Innovación controlada, responsabilidad y confianza institucional	Normas obsoletas, burocracia o vacíos de supervisión	Instituciones y científicos autorizan, auditan y corrigen	Evaluación continua, actores responsables y mecanismos de reclamación

Nota. La tabla presenta escenarios que pueden coexistir según la disciplina, el riesgo y la infraestructura disponible. La evolución hacia mayores niveles de autonomía requiere evaluación independiente, trazabilidad, regulación adaptable y control humano proporcional a las consecuencias científicas.

Los escenarios futuros de incorporación de inteligencia artificial en la investigación científica no representan etapas obligatorias ni excluyentes, sino posibilidades que pueden coexistir según la disciplina, la infraestructura y el nivel de riesgo. La asistencia especializada y la colaboración aumentada ofrecen beneficios inmediatos, mientras los sistemas multiagente, la autonomía supervisada y los laboratorios autónomos exigen controles más estrictos.

Capítulo

6

Implicaciones epistemológicas, profesionales y sociales de la investigación con inteligencia artificial

Rodolfo Marlon Mina Angulo



6.1 Autoridad epistémica y juicio científico en entornos asistidos por inteligencia artificial

La incorporación de inteligencia artificial modifica quién realiza determinadas operaciones intelectuales dentro de una investigación. Los modelos pueden relacionar información, formular hipótesis, comparar alternativas y producir argumentos que anteriormente dependían casi por completo del trabajo humano. Esta ampliación funcional plantea una cuestión epistemológica: realizar una operación cognitiva no equivale a establecer que su resultado constituye conocimiento. Krenn et al. (2022) sostienen que el aporte de la inteligencia artificial a la ciencia debe examinarse también desde su contribución a la comprensión. Una conclusión adquiere valor cuando puede relacionarse razonadamente con evidencia, conceptos y criterios aceptables dentro del campo estudiado.

La aceptación de una afirmación científica requiere razones que permitan distinguirla de una respuesta simplemente plausible. Esta diferencia se vuelve menos evidente cuando los modelos producen textos ordenados, explicaciones aparentemente consistentes y relaciones conceptuales difíciles de distinguir de una argumentación experta. Messeri y Crockett (2024) advierten que la inteligencia artificial puede favorecer ilusiones de comprensión al aumentar la sensación de dominio sobre fenómenos complejos. La fluidez de una respuesta puede influir en la percepción de su calidad, aunque no determine su corrección. El investigador necesita separar capacidad expresiva, plausibilidad argumentativa y conocimiento respaldado por evidencia.

El juicio científico adquiere así una función selectiva. No todas las alternativas generadas poseen igual relevancia, aunque resulten técnicamente posibles o estén formuladas de manera convincente. Figueroa (2025) examina la relación entre inteligencia artificial e investigación desde escenarios de reemplazo, alianza y amenaza, lo que permite reconocer diferentes grados de intervención tecnológica. Frente a una propuesta automatizada, el investigador debe valorar su correspondencia con el problema, los conceptos empleados y las condiciones concretas del fenómeno. La decisión intelectual comienza donde termina la simple generación de opciones, porque elegir exige establecer razones para preferir unas interpretaciones y descartar otras.

Este problema también aparece cuando un modelo obtiene buenos resultados sin ofrecer una representación comprensible del fenómeno analizado. La capacidad predictiva puede

resultar científicamente útil, pero no responde necesariamente a preguntas sobre mecanismos, causas o significado. Schwartz (2022) aborda la discusión acerca de si la inteligencia artificial debe ser interpretable para los seres humanos. En determinados estudios puede bastar una predicción fiable para una finalidad delimitada; en otros, el propósito consiste precisamente en comprender relaciones subyacentes. El criterio de evaluación depende entonces de la pregunta científica y del tipo de conocimiento que se pretende construir mediante el modelo.

Trabajar con sistemas inteligentes exige además reconocer los límites del propio investigador. La familiaridad con una herramienta puede generar confianza operativa sin garantizar capacidad suficiente para valorar sus implicaciones. Chaaban et al. (2024) identifican distintas dimensiones en la preparación de investigadores para utilizar inteligencia artificial, entre ellas componentes técnicos, críticos y éticos. Esta preparación influye en la posibilidad de formular objeciones, reconocer resultados inesperados y reconsiderar decisiones inicialmente aceptadas. El juicio pierde fuerza cuando el investigador únicamente selecciona entre opciones proporcionadas por el sistema y deja de construir criterios propios para determinar su pertinencia.

La colaboración humano-máquina introduce además una influencia recíproca sobre el razonamiento. El investigador formula instrucciones a partir de sus expectativas, mientras las respuestas obtenidas pueden modificar las preguntas siguientes y orientar nuevas decisiones. Shao et al. (2025) muestran que la colaboración con inteligencia artificial puede ampliar alternativas analíticas y favorecer procesos científicos iterativos. Esta dinámica no debe entenderse como una suma independiente de aportes humanos y computacionales. Cada interacción puede alterar la trayectoria intelectual posterior, por lo que resulta necesario reconocer cuándo una idea surge del conocimiento disciplinar previo y cuándo comienza a desarrollarse a partir de alternativas propuestas por el sistema.

La influencia es todavía más compleja cuando varios agentes generan, critican y seleccionan propuestas entre sí. La existencia de debate computacional puede producir una apariencia de evaluación plural, aunque los componentes compartan arquitecturas, información o criterios similares. Tang et al. (2025) señalan riesgos asociados con el aumento de autonomía de sistemas orientados al trabajo científico. Desde una perspectiva epistemológica, el acuerdo entre modelos no debe confundirse con corroboración. Una coincidencia adquiere mayor fuerza cuando procede de aproximaciones suficientemente

independientes y no cuando diferentes componentes reproducen supuestos comunes que permanecen inadvertidos durante el proceso.

También debe considerarse la formación progresiva de una conclusión. En una investigación convencional, el investigador puede reconstruir las razones que lo condujeron desde una observación hasta una determinada interpretación. Cuando interviene inteligencia artificial, algunas conexiones pueden ser sugeridas antes de que la persona haya desarrollado una explicación propia. Díaz (2024) señala que estas tecnologías pueden participar en distintas etapas del trabajo científico. Esta intervención modifica el proceso mediante el cual se forman las convicciones del investigador: una recomendación temprana puede orientar qué relaciones busca posteriormente y favorecer una lectura de los datos consistente con la primera alternativa considerada.

El conocimiento científico tampoco depende exclusivamente del convencimiento individual. Una afirmación se incorpora a un campo cuando puede ser examinada, discutida y eventualmente cuestionada por otros especialistas. Cárdenas (2023) analiza los cambios que la inteligencia artificial puede introducir tanto en investigación como en revisión por pares. La participación de modelos en ambos espacios plantea una dificultad particular: una idea producida con asistencia automatizada podría posteriormente ser valorada mediante sistemas sujetos a regularidades semejantes. La crítica académica conserva su importancia precisamente porque permite confrontar una propuesta con perspectivas que no necesariamente comparten el mismo marco computacional de origen.

La atribución de una conclusión también adquiere un significado distinto cuando la participación tecnológica supera tareas instrumentales. Alonso (2024) examina problemas de integridad académica vinculados con la inteligencia artificial generativa. Desde el plano epistemológico, la cuestión no se limita a identificar quién redactó una determinada frase, sino a establecer quién puede defender intelectualmente el argumento expresado. Firmar un trabajo supone estar en condiciones de explicar sus premisas, responder preguntas sobre sus inferencias y reconocer las razones que sustentan sus conclusiones. Una afirmación que el propio autor no puede defender revela una separación problemática entre producción textual y comprensión.

La utilización de inteligencia artificial tampoco convierte al investigador en un receptor pasivo de resultados inevitables. Comas (2025) plantea interrogantes éticos relacionados con estas herramientas dentro de la investigación, situando su empleo dentro de

decisiones humanas concretas. Elegir un modelo, formular una consulta, conservar una respuesta o modificar una interpretación son acciones que intervienen en la construcción del conocimiento. La participación tecnológica introduce nuevas mediaciones, pero las decisiones relevantes continúan requiriendo una persona capaz de explicar por qué determinada evidencia fue considerada suficiente y por qué una interpretación fue preferida frente a alternativas igualmente posibles.

En este contexto, la autoridad epistémica puede entenderse como la capacidad de sostener razonadamente una afirmación frente al examen científico, más que como una propiedad derivada de haber producido materialmente cada operación del estudio. La inteligencia artificial puede intervenir en la generación de posibilidades, relaciones y argumentos, pero el juicio exige comprender qué se afirma, con qué alcance y bajo qué condiciones podría resultar equivocado. El problema central no consiste en determinar si una máquina puede participar en el razonamiento, sino en evitar que la capacidad de generar inferencias termine sustituyendo la capacidad científica de examinarlas, discutir las y defenderlas.

6.2 Autonomía intelectual y dependencia cognitiva del investigador

La autonomía intelectual del investigador se refiere a su capacidad para formular preguntas, construir criterios de evaluación, interpretar evidencia y sostener decisiones sin depender completamente de una herramienta externa. La inteligencia artificial puede ampliar esas capacidades cuando funciona como apoyo, pero también puede reducirlas si sustituye de manera constante operaciones que antes exigían comprensión directa. Chaaban et al. (2024) muestran que la preparación para trabajar con inteligencia artificial comprende dimensiones técnicas, críticas y éticas. Esa combinación resulta necesaria para que el investigador utilice los sistemas como instrumentos de apoyo sin convertir sus respuestas en referentes incuestionables.

La dependencia cognitiva aparece cuando una herramienta deja de complementar el razonamiento y comienza a condicionar la forma en que se plantean preguntas, se seleccionan fuentes o se interpretan resultados. Messeri y Crockett (2024) advierten que los sistemas de inteligencia artificial pueden generar una sensación de comprensión que supera el conocimiento realmente adquirido. En investigación, este fenómeno puede manifestarse cuando una explicación generada es aceptada porque resulta clara o convincente, aunque el investigador no pueda reconstruir sus fundamentos. La pérdida de

autonomía comienza cuando disminuye la capacidad de evaluar una respuesta sin recurrir nuevamente al mismo sistema que la produjo.

El uso reiterado de asistentes generativos también puede modificar la relación del investigador con la literatura científica. Los sistemas facilitan localizar, resumir y organizar información, pero esa eficiencia puede reducir el contacto directo con las fuentes cuando se emplea de manera indiscriminada. Goyanes y Lopezosa (2024) muestran que la inteligencia artificial puede apoyar tareas de exploración documental en investigación cualitativa y cuantitativa. Sin embargo, la lectura crítica sigue siendo necesaria para reconocer matices, contradicciones y diferencias metodológicas que pueden desaparecer en una síntesis automatizada. Delegar la exploración no debería implicar delegar la comprensión del campo.

Otro ámbito de dependencia se relaciona con el análisis y la programación. Shao et al. (2025) presentan sistemas capaces de automatizar flujos científicos y ampliar alternativas analíticas. Este apoyo puede reducir barreras técnicas, pero también generar una situación en la que el investigador ejecuta procedimientos que no comprende suficientemente. Cuando el código, las transformaciones de datos o las decisiones analíticas se aceptan sin revisión conceptual, disminuye la capacidad para detectar errores y justificar el método utilizado. La autonomía se conserva cuando la asistencia tecnológica mejora la ejecución sin reemplazar la comprensión de las operaciones que sustentan los resultados.

La escritura científica plantea una tensión similar. Los modelos generativos pueden ayudar a organizar argumentos, reformular párrafos y mejorar claridad, pero existe una diferencia entre asistencia lingüística y sustitución del razonamiento. Lopezosa (2023) propone un uso responsable de ChatGPT en comunicación científica, lo que exige mantener control sobre el contenido producido. Si el investigador delega la formulación de explicaciones sin revisar su coherencia conceptual, puede terminar defendiendo argumentos que no construyó ni comprende plenamente. La escritura deja entonces de funcionar como parte del proceso de pensamiento y se convierte en una etapa de aceptación y corrección superficial.

La dependencia también puede aparecer en la toma de decisiones metodológicas. Ruiz (2024) analiza las implicaciones de la inteligencia artificial en distintas fases de la metodología de investigación, donde los sistemas pueden sugerir diseños, técnicas o estrategias de análisis. Estas recomendaciones resultan útiles como alternativas

preliminares, pero no sustituyen la valoración del contexto, los objetivos y las condiciones del estudio. Un investigador cognitivamente dependiente tiende a seleccionar opciones por conveniencia o por frecuencia de recomendación, mientras uno autónomo compara alternativas, identifica supuestos y determina cuál responde mejor al problema investigado.

La formación constituye un mecanismo central para prevenir una dependencia excesiva. Ocampo et al. (2024) relacionan la incorporación de inteligencia artificial con nuevas exigencias en la formación de investigadores, incluyendo capacidades digitales y competencias asociadas con el uso crítico de tecnologías convergentes. El objetivo no consiste únicamente en enseñar a utilizar plataformas, sino en fortalecer habilidades que permitan trabajar también sin ellas. Comprender métodos, leer fuentes originales, interpretar resultados y justificar decisiones conserva la base intelectual necesaria para evaluar posteriormente el apoyo automatizado y reconocer cuándo una recomendación resulta inadecuada.

Tabla 1
Niveles de dependencia cognitiva en el uso de inteligencia artificial durante la investigación

Nivel	Relación con la inteligencia artificial	Manifestación en la investigación	Riesgo principal	Condición para conservar autonomía
Apoyo instrumental	La herramienta ejecuta tareas delimitadas	Organización, búsqueda inicial, formato o cálculo auxiliar	Dependencia operativa menor	Comprender el procedimiento y revisar la salida
Asistencia analítica	El sistema propone métodos, relaciones o interpretaciones	Comparación de alternativas y exploración de resultados	Aceptación acrítica de recomendaciones	Contrastar con criterios metodológicos y disciplinares
Mediación cognitiva	Las respuestas del sistema orientan nuevas preguntas y decisiones	Reformulación de hipótesis, selección de fuentes y lectura de resultados	Sesgo de anclaje y reducción de alternativas	Introducir fuentes independientes y razonamiento propio
Dependencia funcional	El investigador necesita la herramienta para	Programación, análisis, escritura o interpretación	Pérdida de capacidad de auditoría	Mantener competencias básicas y

Nivel	Relación con la inteligencia artificial	Manifestación en la investigación	Riesgo principal	Condición para conservar autonomía
	ejecutar tareas centrales			procedimientos alternativos
Sustitución cognitiva	El sistema produce argumentos y decisiones que el investigador apenas examina	Construcción de explicaciones y conclusiones	Pérdida de comprensión y criterio propio	Recuperar control mediante revisión profunda y justificación independiente

Nota. La tabla distingue distintos grados de dependencia según la función que cumple la inteligencia artificial. El riesgo aumenta cuando el investigador deja de comprender, contrastar o justificar las operaciones y decisiones que el sistema realiza en su nombre.

La dependencia cognitiva no surge únicamente por utilizar inteligencia artificial con frecuencia, sino por la forma en que se distribuyen las tareas intelectuales entre persona y sistema. Figueroa (2025) plantea la relación con estas tecnologías desde escenarios de alianza, reemplazo y amenaza, lo que permite observar que el problema reside en el grado de delegación y no en el uso en sí mismo. Una herramienta puede emplearse intensamente sin eliminar autonomía si el investigador conserva capacidad para cuestionarla, corregirla y trabajar con métodos alternativos cuando sea necesario.

Preservar autonomía intelectual implica mantener una relación activa con el conocimiento producido durante la investigación. La inteligencia artificial puede acelerar procesos, ampliar opciones y reducir tareas repetitivas, pero su utilidad disminuye cuando reemplaza de forma sostenida la comprensión, la argumentación o la capacidad de decidir. La dependencia se vuelve problemática cuando el investigador ya no puede explicar cómo llegó a una conclusión sin recurrir nuevamente al sistema. Mantener criterio propio exige combinar asistencia tecnológica con lectura directa, dominio metodológico, contraste independiente y disposición permanente para rechazar resultados que no puedan ser defendidos razonadamente.

6.3 Mediación algorítmica en la selección de problemas y agendas científicas

La selección de un problema de investigación implica decidir qué fenómenos merecen atención, qué preguntas resultan pertinentes y dónde existen vacíos suficientemente relevantes para justificar un nuevo estudio. La inteligencia artificial interviene cada vez más en esta etapa mediante sistemas capaces de explorar literatura, reconocer recurrencias

temáticas y sugerir áreas de interés. Rodríguez et al. (2025) muestran que las técnicas de inteligencia artificial participan de manera creciente en distintas actividades científicas. Esta incorporación amplía la capacidad de identificar tendencias, aunque también introduce una mediación tecnológica entre el investigador y el campo de conocimiento que pretende examinar.

Los algoritmos de búsqueda y recomendación influyen en qué publicaciones aparecen primero, qué relaciones temáticas se vuelven visibles y qué autores permanecen periféricos dentro de una exploración. Casanova y Martínez (2024), mediante un análisis cienciométrico, evidencian que la producción científica puede representarse a partir de tendencias, conexiones y concentraciones temáticas. Estas representaciones facilitan reconocer estructuras del campo, pero dependen de las bases utilizadas y de los criterios mediante los cuales se organiza la información. Un mapa bibliométrico orienta la observación, aunque no determina por sí mismo qué problema posee mayor importancia científica.

La identificación automatizada de vacíos requiere especial cautela. Un tema con pocas publicaciones puede representar una oportunidad de investigación, pero también reflejar diferencias terminológicas, baja indexación o escasa cobertura documental. Goyanes y Lopezosa (2024) muestran que la inteligencia artificial puede apoyar la exploración de literatura en investigaciones cualitativas y cuantitativas. En este contexto, un vacío sugerido por un sistema debe considerarse provisional hasta comprobar que la ausencia no procede de una búsqueda incompleta. La escasez documental solo adquiere significado cuando se interpreta dentro de la evolución conceptual y metodológica del campo.

Las tendencias de publicación constituyen otro mecanismo capaz de orientar agendas. Los sistemas pueden detectar términos cuya frecuencia aumenta rápidamente y proponerlos como áreas emergentes. Sin embargo, crecimiento documental y relevancia científica no representan necesariamente la misma condición. Casanova y Martínez (2024) muestran cómo el análisis de producción permite observar dinámicas de expansión temática, pero estos indicadores describen comportamiento bibliográfico. Un área puede concentrar numerosos artículos debido a disponibilidad tecnológica, interés editorial o facilidad de acceso a datos. La elección del problema exige valorar también su pertinencia teórica, empírica y disciplinar.

El modo en que se formula una consulta influye igualmente en las alternativas que la inteligencia artificial devuelve al investigador. Zhang et al. (2025) describen la participación de grandes modelos de lenguaje en diferentes etapas del método científico, incluida la generación de hipótesis. Antes de llegar a esa fase, las preguntas introducidas en el sistema ya delimitan conceptos, relaciones y dominios que serán explorados. Una formulación estrecha puede producir un conjunto reducido de posibilidades, mientras otra más amplia modifica el espacio de búsqueda. La mediación algorítmica empieza, por tanto, antes de que aparezca la primera recomendación explícita.

También puede surgir un efecto de retroalimentación cuando los investigadores seleccionan problemas basándose en información generada a partir de la literatura existente. Si los modelos aprenden de campos altamente publicados, estos pueden aparecer repetidamente como áreas relevantes y recibir todavía mayor atención. Messeri y Crockett (2024) advierten que la inteligencia artificial puede producir condiciones que favorecen una percepción limitada de la diversidad de enfoques científicos. En la definición de agendas, esta dinámica puede reforzar temas consolidados mientras cuestiones menos representadas permanecen fuera de los resultados sugeridos, aunque posean importancia en contextos específicos.

La disponibilidad de datos constituye otra fuerza de selección. Los problemas con grandes conjuntos de información estructurada resultan más compatibles con técnicas de aprendizaje automático y pueden recibir mayor atención por su facilidad de procesamiento. Barradas (2023) describe el carácter transformador de la inteligencia artificial en investigación y su capacidad para ampliar el análisis de fenómenos complejos. No obstante, la posibilidad de aplicar un modelo no debería convertirse en criterio principal para definir qué estudiar. Una agenda científica guiada únicamente por datos disponibles corre el riesgo de desplazar preguntas relevantes cuyos fenómenos son difíciles de digitalizar, medir o representar computacionalmente.

La mediación puede afectar además la diversidad disciplinar y regional de las agendas. Choque et al. (2025) relacionan ciencia abierta, inteligencia artificial y colaboración científica, destacando condiciones necesarias para compartir y reutilizar conocimiento. Los sistemas que dependen de repositorios con cobertura desigual pueden privilegiar publicaciones provenientes de instituciones, idiomas o regiones con mayor presencia digital. Como consecuencia, problemas ampliamente estudiados en determinados contextos pueden adquirir visibilidad global, mientras otros permanecen

subrepresentados. La selección responsable exige reconocer que la ausencia en los resultados recuperados no equivale a inexistencia científica ni a falta de relevancia social.

Las instituciones también pueden incorporar inteligencia artificial para reconocer áreas estratégicas, comparar productividad o identificar líneas de investigación con crecimiento acelerado. Guerrero et al. (2025) analizan el impacto de estas tecnologías sobre la producción científica y muestran su presencia dentro de procesos asociados con generación y gestión del conocimiento. Estas aplicaciones pueden aportar información para planificar recursos, pero los indicadores algorítmicos no deberían sustituir la deliberación académica. Una agenda institucional responde también a capacidades humanas, prioridades disciplinares, necesidades del entorno y compromisos científicos que no siempre pueden traducirse adecuadamente en métricas documentales.

Figura 1

Factores que influyen en la mediación algorítmica en la selección de problemas científicos



Nota. Representación de distintos elementos que intervienen en la orientación y delimitación de los temas de investigación.

La generación automatizada de hipótesis puede intensificar este efecto porque conecta selección temática y orientación posterior del estudio. Cuando un sistema propone relaciones consideradas novedosas, algunas pueden convertirse en nuevas preguntas de investigación. Zhang et al. (2025) explican que los modelos pueden colaborar en el

tránsito desde la formulación de hipótesis hasta etapas posteriores del descubrimiento. La utilidad de estas propuestas depende de que el investigador evalúe si representan una contribución real o una recombinación de relaciones frecuentes en los materiales utilizados por el modelo. Novedad generativa y novedad científica deben mantenerse diferenciadas.

Conservar pluralidad en las agendas exige introducir deliberadamente fuentes, enfoques y criterios que no dependan exclusivamente de recomendaciones automatizadas. Rodríguez et al. (2025) identifican una expansión de técnicas de inteligencia artificial dentro de la investigación científica, pero esta expansión ocurre en campos con condiciones y trayectorias diferentes. Contrastar bases documentales, revisar literatura regional, incorporar fuentes especializadas y consultar conocimiento disciplinar permite reducir la influencia de una única representación algorítmica. El investigador debe examinar aquello que el sistema recupera, pero también preguntarse qué perspectivas, problemas y evidencias pueden haber quedado fuera de su campo de visibilidad.

La inteligencia artificial puede ampliar considerablemente el horizonte desde el cual se identifican problemas científicos, aunque también interviene en la construcción de aquello que parece relevante, novedoso o investigable. Por esta razón, las agendas no deberían derivarse automáticamente de tendencias, frecuencias o recomendaciones generadas por modelos. La selección de un problema continúa exigiendo valorar su fundamento conceptual, la existencia de evidencia insuficiente, su relación con preguntas disciplinares y las condiciones concretas donde será estudiado. La mediación algorítmica resulta útil cuando amplía el campo de observación, pero se vuelve restrictiva cuando termina definiendo sus límites.

6.4 Homogeneización del conocimiento y preservación de la diversidad científica

La expansión de la inteligencia artificial en la investigación introduce la posibilidad de que distintos equipos utilicen modelos similares para buscar información, organizar antecedentes, formular hipótesis y redactar interpretaciones. Esta convergencia puede aumentar la eficiencia, pero también reducir la variedad de enfoques cuando los sistemas reproducen patrones dominantes presentes en sus datos de entrenamiento. Messeri y Crockett (2024) advierten que determinadas formas de asistencia automatizada pueden estrechar la percepción de alternativas disponibles. La homogeneización aparece cuando

múltiples investigaciones terminan recurriendo a estructuras conceptuales semejantes sin examinar suficientemente perspectivas diferentes.

La diversidad científica depende de que existan múltiples formas de formular problemas, seleccionar métodos e interpretar evidencia. Esta pluralidad permite confrontar explicaciones y evitar que una única representación del fenómeno adquiera carácter dominante por mera repetición. Krenn et al. (2022) señalan que la inteligencia artificial puede contribuir a la comprensión científica cuando sus resultados se articulan con conocimiento teórico y disciplinar. Desde esta perspectiva, conservar diversidad implica utilizar modelos como instrumentos de exploración, no como fuentes únicas para definir conceptos, argumentos o trayectorias de investigación.

El riesgo de uniformidad aumenta cuando los investigadores dependen de las mismas bases documentales, asistentes generativos o sistemas de recomendación. Los algoritmos tienden a priorizar contenidos con mayor presencia, conectividad o compatibilidad con los patrones previamente aprendidos. Rodríguez et al. (2025) muestran que las técnicas de inteligencia artificial se extienden de manera creciente dentro de la actividad científica. Si esta expansión ocurre sobre infraestructuras similares, determinados enfoques pueden adquirir mayor visibilidad mientras otros permanecen periféricos, especialmente cuando proceden de tradiciones académicas con menor representación digital.

El idioma constituye una dimensión especialmente sensible. Gran parte de la producción científica internacional se concentra en inglés, lo que puede influir en los materiales recuperados y en las respuestas generadas por sistemas entrenados con corpus desiguales. Choque et al. (2025) relacionan la ciencia abierta con la necesidad de ampliar colaboración y acceso al conocimiento. Preservar diversidad lingüística no supone únicamente traducir contenidos, sino mantener conceptos, categorías y debates desarrollados en contextos específicos. Cuando la mediación algorítmica favorece un vocabulario dominante, ciertas formas de conocimiento pueden quedar subordinadas o simplificadas.

Las diferencias disciplinares también pueden debilitarse cuando herramientas generalistas aplican patrones similares a problemas científicos heterogéneos. Una estrategia adecuada en ciencias experimentales no necesariamente conserva su validez en educación, ciencias sociales o investigación cualitativa. Goyanes y Lopezosa (2024) muestran que el uso de inteligencia artificial en ciencias sociales requiere atender particularidades metodológicas

de los enfoques cuantitativos y cualitativos. La diversidad científica se preserva cuando cada disciplina mantiene criterios propios para definir evidencia, interpretar resultados y determinar qué tipos de explicación son pertinentes para sus objetos de estudio.

Otro problema surge cuando la automatización favorece formatos estandarizados de argumentación. Los modelos generativos pueden producir introducciones, discusiones o conclusiones con estructuras recurrentes, especialmente si reciben instrucciones semejantes. Lopezosa (2023) examina el uso responsable de ChatGPT en comunicación científica y plantea la necesidad de mantener control crítico sobre los contenidos producidos. La estandarización formal puede facilitar claridad, pero se vuelve problemática cuando afecta el desarrollo conceptual. Una investigación puede cumplir correctamente una estructura textual y, al mismo tiempo, ofrecer escasa diferenciación teórica respecto de trabajos contruidos mediante patrones similares.

La preservación de diversidad requiere introducir deliberadamente fuentes y perspectivas que los sistemas no recuperan de forma prioritaria. Esto incluye literatura regional, publicaciones en distintos idiomas, trabajos interdisciplinarios y resultados que contradicen las tendencias dominantes. Casanova y Martínez (2024) muestran que los análisis cuenciométricos permiten identificar concentraciones y relaciones dentro de la producción científica, pero también evidencian desigualdades de visibilidad entre temas. Examinar aquello que aparece con menor frecuencia puede ser tan importante como identificar los núcleos más consolidados, especialmente cuando se pretende construir preguntas nuevas o revisar supuestos establecidos.

Tabla 2
Factores de homogeneización y mecanismos para preservar la diversidad científica

Factor de homogeneización	Manifestación posible	Consecuencia científica	Mecanismo de preservación
Uso de modelos similares	Respuestas y estructuras argumentativas recurrentes	Reducción de alternativas conceptuales	Comparar resultados con enfoques no automatizados
Dependencia de bases dominantes	Mayor visibilidad de determinados autores y regiones	Exclusión de literatura periférica	Incorporar repositorios regionales y fuentes especializadas

Factor de homogeneización	Manifestación posible	Consecuencia científica	Mecanismo de preservación
Predominio lingüístico	Priorización de literatura y conceptos en idiomas mayoritarios	Pérdida de matices culturales y conceptuales	Consultar producción científica multilingüe
Estandarización metodológica	Aplicación repetida de procedimientos similares	Menor adaptación al objeto de estudio	Mantener criterios disciplinares específicos
Síntesis automatizada	Eliminación de contradicciones y diferencias entre estudios	Simplificación del debate científico	Revisar textos originales y resultados divergentes
Recomendación algorítmica	Repetición de temas y enfoques populares	Concentración de agendas de investigación	Explorar deliberadamente líneas minoritarias o emergentes
Automatización de escritura	Uniformidad en argumentos y organización textual	Disminución de diferenciación intelectual	Revisar estructura, razonamiento y lenguaje desde criterios propios

Nota. La tabla identifica mecanismos mediante los cuales el uso extendido de inteligencia artificial puede favorecer convergencias excesivas en la producción científica y plantea acciones orientadas a conservar pluralidad conceptual, metodológica, lingüística y disciplinar.

La diversidad también depende de la capacidad institucional para sostener líneas de investigación que no responden inmediatamente a tendencias tecnológicas dominantes. Ocampo et al. (2024) relacionan la formación de investigadores con transformaciones producidas por tecnologías convergentes y nuevas competencias científicas. En este escenario, las instituciones deben evitar que la adopción de inteligencia artificial transforme la innovación tecnológica en criterio único de pertinencia. Mantener espacios para investigación teórica, interpretativa, local o metodológicamente minoritaria permite conservar un sistema científico más amplio que el definido por las posibilidades de automatización disponibles.

La colaboración entre investigadores y sistemas puede ser productiva si se utiliza para ampliar alternativas en lugar de reducirlas. Shao et al. (2025) muestran que la interacción humano-IA puede favorecer procesos iterativos y exploración de diferentes opciones analíticas. Para que esa colaboración preserve diversidad, las propuestas generadas deben ser contrastadas con marcos conceptuales alternativos, datos procedentes de contextos

distintos y criterios elaborados por especialistas. La existencia de múltiples respuestas computacionales no garantiza pluralidad si todas derivan de supuestos similares; la diversidad requiere diferencias reales en fuentes, métodos y perspectivas.

La homogeneización del conocimiento no constituye una consecuencia inevitable del uso de inteligencia artificial, sino un riesgo asociado con formas repetitivas y poco críticas de incorporación. Preservar diversidad científica exige mantener variedad de preguntas, tradiciones disciplinares, idiomas, contextos y procedimientos de validación. Los sistemas inteligentes pueden ayudar a conectar perspectivas y localizar relaciones poco evidentes, pero su utilidad depende de que no reduzcan el espacio intelectual a los patrones más frecuentes. Una ciencia asistida por inteligencia artificial conserva mayor capacidad explicativa cuando utiliza la tecnología para ampliar posibilidades y no para uniformarlas.

6.5 Asimetrías de acceso y concentración de capacidades científicas

La incorporación de inteligencia artificial en la investigación depende de recursos que no están distribuidos de manera uniforme entre instituciones, disciplinas y regiones. El acceso a capacidad de cómputo, repositorios especializados, plataformas de pago, infraestructura de datos y personal técnico condiciona qué equipos pueden utilizar sistemas avanzados de forma sostenida. Rodríguez et al. (2025) muestran que la inteligencia artificial ocupa un lugar creciente dentro de la actividad científica, aunque su aprovechamiento ocurre en contextos institucionales diferentes. Esta desigualdad puede traducirse en capacidades distintas para procesar información, automatizar procedimientos y desarrollar proyectos de mayor complejidad tecnológica.

La infraestructura constituye una de las principales fuentes de asimetría. Los proyectos que requieren procesamiento intensivo, almacenamiento extenso o integración de múltiples fuentes necesitan recursos tecnológicos que no siempre están disponibles en universidades pequeñas o centros con financiamiento limitado. Tom et al. (2024) describen los laboratorios autónomos como entornos que combinan instrumentación, automatización, aprendizaje automático y planificación computacional. Estas arquitecturas muestran que determinados avances científicos dependen tanto del conocimiento especializado como de inversiones materiales difíciles de reproducir en instituciones con menor capacidad presupuestaria.

Las diferencias también aparecen en el acceso a modelos y servicios digitales. Algunas plataformas ofrecen versiones gratuitas con funciones limitadas, mientras las herramientas de mayor capacidad pueden exigir suscripciones, consumo de servicios en la nube o contratos institucionales. Guerrero et al. (2025) analizan el impacto de la inteligencia artificial sobre la producción científica y su incorporación en diferentes tareas investigativas. Cuando el acceso a sistemas más potentes depende de recursos económicos, la capacidad de explorar grandes corpus, ejecutar análisis complejos o automatizar etapas puede concentrarse en grupos con mejores condiciones financieras.

La disponibilidad y calidad de los datos introduce otra desigualdad. Algunos campos cuentan con repositorios amplios, bien documentados y reutilizables, mientras otros dependen de registros fragmentados, poco digitalizados o de acceso restringido. Choque et al. (2025) vinculan la ciencia abierta con la colaboración y circulación del conocimiento, destacando la importancia de disponer de recursos que puedan ser compartidos y reutilizados. Sin condiciones semejantes de acceso, ciertos equipos pueden entrenar, contrastar o adaptar modelos con mayor facilidad, mientras otros enfrentan limitaciones antes incluso de iniciar el análisis computacional.

El dominio lingüístico y la visibilidad internacional también intervienen en la distribución de capacidades. La concentración de publicaciones, documentación técnica y herramientas en determinados idiomas puede favorecer a investigadores con mayor acceso a esos circuitos. Casanova y Martínez (2024) muestran, mediante análisis de producción científica, que los campos de investigación presentan concentraciones temáticas y documentales que permiten observar diferencias de visibilidad. En entornos asistidos por inteligencia artificial, estas desigualdades pueden ampliarse cuando los sistemas recuperan con mayor facilidad literatura procedente de fuentes dominantes y reducen la presencia de trabajos regionales o menos indexados.

Las capacidades humanas constituyen una dimensión adicional. Disponer de herramientas avanzadas no garantiza un aprovechamiento científico adecuado si faltan competencias para evaluar modelos, revisar código, gestionar datos o interpretar resultados. Chaaban et al. (2024) identifican que la preparación para trabajar con inteligencia artificial incluye componentes técnicos, críticos y éticos. Las instituciones con programas permanentes de formación pueden integrar estas tecnologías con mayor autonomía, mientras aquellas que dependen de iniciativas individuales corren mayor riesgo de utilizar servicios externos sin suficiente capacidad para comprender sus límites o revisar sus decisiones.

La concentración tecnológica también puede generar dependencia respecto de proveedores privados. Cuando una investigación utiliza modelos, infraestructuras o servicios cuya arquitectura, actualización o condiciones de acceso son controladas externamente, parte de la continuidad del proyecto queda sujeta a decisiones ajenas al equipo científico. Van et al. (2024) destacan la importancia de organizar y documentar los componentes de la investigación computacional para fortalecer reproducibilidad y transparencia. Esta necesidad adquiere mayor relevancia cuando los recursos utilizados no pueden conservarse, replicarse o mantenerse disponibles en condiciones equivalentes para otros grupos.

Tabla 3
Asimetrías en el acceso a capacidades científicas asociadas con inteligencia artificial

Dimensión	Condición de mayor capacidad	Condición de menor capacidad	Efecto posible sobre la investigación	Medida de equilibrio
Infraestructura	Cómputo avanzado, almacenamiento y automatización	Equipos limitados y baja capacidad de procesamiento	Diferencias en escala y complejidad de los proyectos	Infraestructura compartida y servicios institucionales
Acceso a modelos	Plataformas avanzadas y servicios especializados	Versiones restringidas o acceso discontinuo	Menor capacidad para comparar o adaptar herramientas	Licencias institucionales y alternativas abiertas
Datos	Repositorios amplios, estructurados y documentados	Datos fragmentados, escasos o poco digitalizados	Limitaciones para entrenamiento y validación	Repositorios interoperables y políticas de apertura
Formación	Equipos con competencias técnicas y metodológicas	Dependencia de conocimiento externo	Menor capacidad de revisión y auditoría	Programas continuos de formación
Visibilidad lingüística	Acceso a literatura y herramientas dominantes	Menor representación regional o idiomática	Sesgo en fuentes y problemas considerados	Inclusión de repositorios y producción multilingüe
Dependencia tecnológica	Control institucional de recursos y versiones	Uso de servicios externos sin alternativas	Vulnerabilidad ante cambios de acceso o condiciones	Documentación, portabilidad y diversificación de herramientas

Nota. La tabla sintetiza las principales desigualdades que condicionan el acceso y uso de inteligencia artificial en la investigación.

Las asimetrías de acceso pueden influir en la distribución de oportunidades para publicar, colaborar y participar en proyectos de alta intensidad tecnológica. Ocampo et al. (2024) relacionan las tecnologías convergentes con nuevas exigencias para la formación de investigadores, lo que evidencia que la transformación científica también depende de condiciones institucionales de preparación. Si únicamente determinados centros pueden sostener equipos especializados, infraestructura avanzada y acceso continuo a modelos, la brecha no se limita al uso de herramientas: puede afectar la posibilidad de definir agendas, producir evidencia y competir por reconocimiento académico.

Reducir estas diferencias exige ampliar las condiciones materiales y formativas que permiten utilizar inteligencia artificial con independencia suficiente. La ciencia abierta, los recursos compartidos, la documentación técnica y las alternativas interoperables pueden disminuir parte de la concentración, aunque no eliminan por sí mismas las desigualdades económicas e institucionales. La capacidad científica debe evaluarse considerando quién controla los datos, los modelos y la infraestructura necesaria para producir resultados. Cuando esos recursos se concentran de manera sostenida, también puede concentrarse la posibilidad de decidir qué problemas se investigan, qué métodos se desarrollan y qué conocimientos alcanzan mayor visibilidad.

6.6 Transformación del trabajo investigador y de la experticia científica

La incorporación de inteligencia artificial está modificando la distribución de tareas dentro de la investigación científica. Actividades como búsqueda documental, programación, análisis preliminar, visualización y redacción pueden desarrollarse ahora con distintos grados de asistencia automatizada. Guerrero et al. (2025) analizan el impacto de estas tecnologías sobre la producción científica y muestran su presencia creciente en diferentes fases del trabajo académico. Este cambio no elimina la actividad del investigador, pero transforma qué operaciones realiza directamente, cuáles supervisa y en qué momentos debe intervenir con mayor profundidad conceptual o metodológica.

La transformación también afecta la forma en que se entiende la experticia científica. Tradicionalmente, el dominio de un campo se asociaba con conocimiento acumulado, experiencia metodológica y capacidad para resolver problemas especializados. En entornos asistidos por inteligencia artificial, parte de esas operaciones puede externalizarse hacia sistemas que recuperan información, generan código o proponen interpretaciones. Figueroa (2025) examina esta relación desde escenarios de alianza y

reemplazo. La experticia deja de depender únicamente de ejecutar procedimientos y pasa a incluir la capacidad para decidir cuándo una herramienta resulta pertinente y cómo debe integrarse al proceso.

El investigador comienza a asumir con mayor frecuencia funciones de coordinación. En lugar de ejecutar cada tarea de forma manual, puede diseñar flujos, conectar herramientas y revisar productos generados por distintos sistemas. Shao et al. (2025) muestran que la colaboración humano-IA puede organizar procesos iterativos y ampliar alternativas analíticas. Esta reorganización modifica el trabajo cotidiano porque desplaza parte del esfuerzo desde la producción directa hacia la selección, supervisión y articulación de resultados. Coordinar adecuadamente exige conocer suficientemente cada fase para detectar incompatibilidades entre datos, métodos y conclusiones.

La programación constituye un ejemplo claro de esta transición. Los modelos pueden generar scripts, explicar errores y sugerir modificaciones, reduciendo barreras para investigadores con formación computacional limitada. Sin embargo, la automatización no elimina la necesidad de comprender qué operaciones ejecuta el código. Vimos et al. (2024) analizan el uso de inteligencia artificial en procesos de investigación por parte de docentes universitarios y muestran su incorporación en distintas actividades científicas. La experticia técnica se transforma cuando el valor profesional reside menos en escribir cada línea y más en evaluar si el procedimiento representa correctamente el diseño metodológico.

Algo similar ocurre con el análisis de datos. Los sistemas pueden recomendar pruebas, reconocer patrones y producir interpretaciones preliminares, pero estas funciones no sustituyen el conocimiento necesario para determinar si un análisis responde al problema planteado. Ruiz (2024) aborda las implicaciones de la inteligencia artificial en la metodología de investigación y muestra que su intervención alcanza decisiones que antes dependían directamente del investigador. La transformación profesional exige diferenciar entre capacidad para obtener un resultado y capacidad para comprender sus supuestos, límites y consecuencias dentro de un diseño científico concreto.

La búsqueda y gestión de literatura también están modificando tareas tradicionalmente asociadas con la experiencia académica. Los investigadores pueden explorar colecciones extensas y obtener síntesis en periodos reducidos, lo que disminuye parte del trabajo manual de clasificación. Goyanes y Lopezosa (2024) examinan aplicaciones de

inteligencia artificial en métodos de investigación y muestran su utilidad en tareas documentales. Sin embargo, la experiencia disciplinar continúa siendo necesaria para distinguir contribuciones centrales, debates secundarios y diferencias conceptuales que una organización automatizada puede presentar como equivalentes.

La escritura científica representa otra área donde cambian las fronteras de la experticia. Los modelos generativos pueden apoyar estructuración, reformulación y corrección lingüística, reduciendo el tiempo dedicado a determinadas tareas editoriales. Lopezosa (2023) plantea que la inteligencia artificial puede ser útil en comunicación científica siempre que se emplee responsablemente. La competencia del investigador se desplaza entonces hacia la capacidad de preservar precisión conceptual, coherencia argumentativa y correspondencia entre texto y evidencia. Redactar con asistencia no disminuye esas exigencias; puede aumentarlas cuando el contenido generado requiere una revisión más rigurosa.

La automatización también modifica las relaciones dentro de los equipos de investigación. Algunas tareas que antes exigían especialistas separados pueden ser parcialmente ejecutadas mediante herramientas accesibles a otros miembros. Esto puede favorecer mayor integración entre funciones, aunque también genera el riesgo de subestimar la complejidad profesional de determinadas áreas. Chaaban et al. (2024) muestran que la preparación para trabajar con inteligencia artificial incorpora capacidades diversas y no exclusivamente instrumentales. El acceso a una herramienta especializada no equivale a poseer la formación necesaria para interpretar correctamente sus resultados ni para asumir responsabilidades propias de otra disciplina.

Las nuevas funciones incrementan la importancia de la capacidad de auditoría. Cuando una tarea es delegada a un sistema, el investigador debe poder examinar suficiente información para determinar si fue ejecutada de acuerdo con los objetivos científicos. Van et al. (2024) desarrollan una propuesta orientada a mejorar reproducibilidad y transparencia en investigación computacional mediante una organización sistemática de los componentes del estudio. En este contexto, documentar procedimientos permite que la experticia no dependa únicamente de interactuar con una plataforma, sino de comprender cómo se conectan datos, decisiones, análisis y resultados.

La especialización científica tampoco desaparece frente a modelos de propósito general. Aunque estos sistemas pueden intervenir en múltiples áreas, sus respuestas necesitan ser

interpretadas según conceptos, métodos y criterios específicos de cada disciplina. Krenn et al. (2022) sostienen que la inteligencia artificial debe contribuir a la comprensión científica y no limitarse al rendimiento predictivo. Esta distinción refuerza el valor del conocimiento especializado, porque determinar qué significa un resultado requiere reconocer mecanismos, condiciones de aplicación y relaciones con teorías previas. La tecnología amplía operaciones posibles, pero no uniformiza los criterios de validez entre campos.

Figura 2
Transformación del trabajo investigador



Nota. Esquema del cambio progresivo en las funciones del investigador conforme aumenta la integración de herramientas de inteligencia artificial.

El crecimiento de sistemas multiagente introduce una transformación adicional. Algunas arquitecturas pueden distribuir funciones entre agentes que generan hipótesis, revisan propuestas, ejecutan análisis y comparan resultados. Ghareeb et al. (2026) presentan un sistema multiagente orientado a automatizar etapas del descubrimiento científico. Estas configuraciones desplazan parte del trabajo del investigador hacia la definición de objetivos, roles y condiciones de intervención. La experticia se expresa entonces en la

capacidad para diseñar un entorno de trabajo donde las funciones automatizadas permanezcan alineadas con preguntas científicas concretas.

La formación profesional necesita responder a estos cambios sin reducirse al aprendizaje de herramientas específicas. Ocampo et al. (2024) relacionan las tecnologías convergentes y la inteligencia artificial con transformaciones en la formación de investigadores. Las plataformas pueden cambiar rápidamente, mientras habilidades como interpretación, evaluación metodológica, lectura crítica y comprensión disciplinar conservan mayor estabilidad. Por ello, la preparación científica debe combinar competencias tecnológicas con fundamentos que permitan adaptarse a sistemas nuevos sin depender de instrucciones operativas particulares. La actualización profesional adquiere sentido cuando fortalece capacidad de decisión y no solo familiaridad con interfaces.

La transformación del trabajo investigador conduce a una experticia más orientada a integrar, evaluar y justificar que a ejecutar manualmente todas las operaciones del proceso. Esta modificación no implica que las habilidades tradicionales pierdan valor, porque siguen siendo necesarias para comprender aquello que se delega y reconocer errores. El investigador que trabaja con inteligencia artificial necesita combinar conocimiento disciplinar, criterio metodológico y capacidad para coordinar herramientas heterogéneas. La experticia científica se redefine, por tanto, alrededor de la posibilidad de mantener control intelectual sobre procesos cada vez más distribuidos entre personas, modelos y sistemas automatizados.

6.7 Confianza y legitimidad del conocimiento científico mediado por inteligencia artificial

La confianza en el conocimiento científico depende de que las afirmaciones puedan ser examinadas, justificadas y contrastadas por quienes participan en su producción y evaluación. Cuando interviene inteligencia artificial, esa confianza incorpora una condición adicional: comprender hasta qué punto los resultados dependen de sistemas que procesan información mediante procedimientos parcialmente opacos. Barberis et al. (2024) destacan la importancia de robustez y reproducibilidad en aplicaciones científicas de inteligencia artificial. La confianza no puede descansar únicamente en el rendimiento del modelo, sino en la consistencia de los resultados bajo condiciones comparables y en la posibilidad de someterlos a comprobación independiente.

La legitimidad científica se construye mediante procedimientos reconocibles para una comunidad especializada. Un resultado puede ser novedoso o técnicamente sofisticado y, aun así, carecer de aceptación si no existe una base suficiente para evaluar cómo fue producido. Krenn et al. (2022) sostienen que la inteligencia artificial debe contribuir a la comprensión científica y no limitarse a generar predicciones eficaces. Esta distinción resulta relevante porque la legitimidad de una conclusión no proviene del prestigio tecnológico del sistema utilizado, sino de su correspondencia con evidencia, métodos adecuados y explicaciones compatibles con el conocimiento disciplinar disponible.

La apariencia de seguridad constituye un problema particular en los modelos generativos. Estos sistemas pueden producir argumentos fluidos incluso cuando la información presentada contiene errores o relaciones insuficientemente justificadas. Athaluri et al. (2023) documentaron referencias inexistentes generadas por ChatGPT en escritura científica, mostrando que la coherencia textual no garantiza exactitud bibliográfica. Situaciones de este tipo afectan la confianza porque dificultan diferenciar rápidamente entre contenido verificable y material fabricado. La credibilidad del conocimiento mediado por inteligencia artificial exige que las fuentes, datos y afirmaciones relevantes puedan comprobarse fuera del entorno que generó la respuesta.

La confianza también está relacionada con la estabilidad de los procedimientos computacionales. Cambios de versión, parámetros, datos de entrada o configuraciones pueden alterar los resultados obtenidos con una misma herramienta. Van et al. (2024) proponen mediante ENCORE una organización orientada a mejorar la reproducibilidad y transparencia de la investigación computacional. Registrar los elementos que intervienen en cada análisis permite reconstruir condiciones de producción y comparar resultados entre momentos diferentes. Esta documentación adquiere especial importancia cuando los sistemas evolucionan con rapidez y una investigación no puede depender de que una plataforma conserve indefinidamente las mismas características.

La legitimidad disminuye cuando el papel de la inteligencia artificial permanece ambiguo dentro de una investigación. No produce el mismo efecto utilizar un sistema para corregir estilo que emplearlo para seleccionar datos, generar código o interpretar resultados. Gimeno y Trigo (2024) examinan el papel de estas tecnologías en la publicación científica y las cuestiones que introducen en el proceso editorial. Diferenciar las funciones automatizadas permite valorar adecuadamente su incidencia sobre el producto final. Una declaración genérica de uso ofrece poca información si no permite reconocer qué

componentes del razonamiento o del procedimiento estuvieron condicionados por el sistema.

La revisión por pares constituye uno de los mecanismos mediante los cuales la comunidad científica examina la credibilidad de nuevas contribuciones. Cárdenas (2023) analiza escenarios relacionados con inteligencia artificial, investigación y evaluación por pares, señalando transformaciones que pueden afectar este proceso. La utilización de sistemas automatizados por autores y revisores puede agilizar determinadas tareas, pero también producir coincidencias derivadas de modelos semejantes. La legitimidad requiere conservar una evaluación suficientemente independiente, donde las conclusiones puedan ser cuestionadas desde criterios disciplinares distintos de aquellos utilizados durante su generación.

El carácter colectivo de la confianza científica implica además reconocer errores y corregirlos cuando aparecen nuevas evidencias. La inteligencia artificial puede incrementar la velocidad con que se producen análisis y manuscritos, pero esa aceleración no debería reducir los tiempos necesarios para comprobar resultados. Alonso (2024) aborda la integridad académica en el contexto de la inteligencia artificial generativa y destaca la necesidad de preservar prácticas responsables. La confianza se fortalece cuando existen mecanismos para identificar fallas, rectificar información y explicar qué modificaciones fueron realizadas, en lugar de mantener conclusiones únicamente porque ya han sido difundidas o publicadas.

La representatividad de los datos influye directamente sobre la legitimidad de resultados asistidos por modelos. Un sistema puede presentar buen desempeño general y ofrecer conclusiones deficientes para grupos, contextos o regiones escasamente representados. Truhn et al. (2024) muestran la importancia de integrar información diversa en modelos fundacionales multimodales aplicados a oncología de precisión. La extrapolación de resultados exige conocer las poblaciones y condiciones a partir de las cuales fueron contruidos. Una conclusión pierde fuerza cuando pretende aplicarse a contextos que no estuvieron suficientemente presentes en los datos utilizados para producirla.

La confianza pública incorpora exigencias diferentes de las que operan exclusivamente dentro de comunidades académicas. Cuando los resultados asistidos por inteligencia artificial influyen en salud, educación, políticas institucionales u otros ámbitos sociales, las personas afectadas necesitan razones comprensibles para aceptar determinadas

decisiones. Comas (2025) plantea interrogantes éticos sobre el empleo de estas herramientas durante una investigación. La legitimidad social depende de que los beneficios esperados no se separen de obligaciones relacionadas con privacidad, consentimiento y uso adecuado de información. La capacidad técnica del sistema no elimina la necesidad de considerar a quienes reciben sus consecuencias.

La apertura científica puede fortalecer la confianza cuando permite acceder a publicaciones, datos, código y protocolos utilizados en un estudio. Choque et al. (2025) relacionan ciencia abierta e inteligencia artificial con formas ampliadas de colaboración y circulación del conocimiento. Sin embargo, apertura no significa publicar indiscriminadamente toda la información disponible. Algunos conjuntos contienen datos personales o materiales sujetos a restricciones justificadas. La legitimidad surge de equilibrar posibilidad de revisión con protección de información sensible, estableciendo condiciones de acceso que permitan comprobar resultados sin vulnerar derechos o comprometer la seguridad de participantes.

Los sistemas con mayor autonomía exigen criterios de confianza proporcionales a las decisiones que pueden ejecutar. Tang et al. (2025) advierten sobre los riesgos asociados con científicos artificiales y proponen priorizar salvaguardas frente al incremento de autonomía. Un asistente que organiza bibliografía no requiere el mismo nivel de control que un agente capaz de ejecutar código o intervenir en experimentos. La aceptación científica de estos sistemas depende de que existan límites adecuados a su capacidad de actuación. A medida que aumenta la autonomía, también aumenta la necesidad de comprobar que sus operaciones permanecen dentro de condiciones previamente definidas.

La legitimidad tampoco puede establecerse exclusivamente mediante indicadores cuantitativos de rendimiento. Precisión, velocidad o capacidad de procesamiento permiten comparar modelos, pero no determinan por sí mismas la pertinencia científica de sus conclusiones. Messeri y Crockett (2024) advierten que la inteligencia artificial puede favorecer ilusiones de comprensión cuando amplía la productividad sin mejorar necesariamente la diversidad o profundidad del conocimiento. Evaluar un sistema exige considerar qué preguntas permite responder, qué aspectos deja fuera y cómo modifica las interpretaciones disponibles. Una tecnología eficiente puede seguir siendo insuficiente para determinados objetivos científicos.

La confianza en conocimiento mediado por inteligencia artificial se sostiene cuando la participación tecnológica permanece compatible con las condiciones que permiten someter una afirmación a crítica. La comunidad científica necesita conocer qué evidencia respalda los resultados, qué limitaciones acompañan al modelo y bajo qué circunstancias una conclusión debería modificarse o rechazarse. La legitimidad no proviene de excluir la inteligencia artificial del proceso investigador, sino de impedir que su complejidad tecnológica dificulte la evaluación del conocimiento producido. Cuando las afirmaciones conservan posibilidad de contraste, discusión y revisión, la mediación algorítmica puede integrarse sin debilitar los fundamentos de la confianza científica.

Bibliografía

- Abolhasani, M., y Kumacheva, E. (2023). The rise of self-driving labs in chemical and materials sciences. *Nature Synthesis*, 2, 483-492. <https://doi.org/10.1038/s44160-022-00231-0>
- Alonso, J. (2024). Integridad académica en la era de la inteligencia artificial generativa. *Desiderata*, 24, 72-81. <https://dialnet.unirioja.es/descarga/articulo/9702543.pdf>
- 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), e37432. <https://doi.org/10.7759/cureus.37432>
- Barberis, A., Aerts, H. J., y Buffa, F. M. (2024). Robustness and reproducibility for AI learning in biomedical sciences. *Scientific Reports*, 14(1), 1933. <https://doi.org/10.1038/s41598-024-51381-4>
- Barradas, G. J. (2023). Inteligencia artificial como elemento transformador de la investigación científica. *Entrelineas*, 2(1), 113-122. <https://doi.org/10.56368/Entrelineas213>
- Barrios, I. (2023). Inteligencia artificial y redacción científica: aspectos éticos en el uso de las nuevas tecnologías. *Medicina Clínica y Social*, 7(2), 46-47. <https://doi.org/10.52379/mcs.v7i2.278>
- Behrend, T., y Landers, R. (2025). Participant interactions with artificial intelligence: using large language models to generate research materials for surveys and experiments. *Journal of Business and Psychology*, 40, 1275-1297. <https://doi.org/10.1007/s10869-025-10035-6>
- Bennis, I., y Mouwafaq, S. (2025). Advancing AI-driven thematic analysis in qualitative research: a comparative study of nine generative models on Cutaneous Leishmaniasis data. *BMC Medical Informatics and Decision Making*, 25(1), 124. <https://doi.org/10.1186/s12911-025-02961-5>
- Berrios, Z. R., y Márquez, M. L. (2025). Agente de inteligencia artificial generativa en investigación científica. *Un análisis explicativo del aprendizaje en el aula. RIED*.

- Revista Iberoamericana de Educación a Distancia*, 28(2), 39-55.
<https://doi.org/10.5944/ried.28.2.43545>
- Boiko, D., MacKnight, R., Kline, B., y Gomes, G. (2023). Autonomous chemical research with large language models. *Nature*, 624, 570-578.
<https://doi.org/10.1038/s41586-023-06792-0>
- Cano, F., Schonhaut, L., Harris, P. R., Millán, T., Mena, P., Lizama, M., . . . Weisstaub, G. (2023). Andes Pediatría y publicaciones científicas en la era de la inteligencia artificial. *Andes Pediatría*, 94(6), 667-671.
<https://doi.org/10.32641/andespediatr.v94i6.4974>
- Cárdenas, J. (2023). Inteligencia artificial, investigación y revisión por pares: escenarios futuros y estrategias de acción. *Revista Española de Sociología*, 32(4), a184, 1-15. <https://doi.org/10.22325/fes/res.2023.184>
- Casanova, P. A., y Martínez, D. M. (2024). Producción científica sobre inteligencia artificial y educación: un análisis cuantitativo. *Hachetetepe. Revista Científica de Educación y Comunicación*(28), 1-23.
<https://doi.org/10.25267/Hachetetepe.2024.i28.1102>
- Cedeño, M. J., Maitta, R. I., Vélez, Z. M., y Palomeque, Z. J. (2024). Investigación universitaria con inteligencia artificial. *Revista Venezolana de Gerencia*, 29(106), 817-830. <https://doi.org/10.52080/rvgluz.29.106.23>
- Chaaban, Y., Qadhi, S., Chen, J., y Du, X. (2024). Understanding researchers' AI readiness in a higher education context: Q methodology research. *Education Sciences*, 14(7), 709. <https://doi.org/10.3390/educsci14070709>
- Choque, G. J., Mamani, F. J., Puita, C. G., Castro, A. D., y Choque, M. J. (2025). Ciencia abierta y colaborativa en la era de la inteligencia artificial. *Revista de Divulgação Científica Sena Aires*, 6(1), 2162-2174. <https://doi.org/10.61616/rvdc.v6i1.505>
- Cisneros, C. A., Guevara, G. A., Urdánigo, C. J., y Garcés, B. J. (2022). Técnicas e instrumentos para la recolección de datos que apoyan a la investigación científica en tiempo de pandemia. *Dominio de las Ciencias*, 8(1), 1165-1185.
<https://dialnet.unirioja.es/servlet/articulo?codigo=8383508>

- Comas, R. R. (2025). ¿Es ético el uso de herramientas de inteligencia artificial en el desarrollo de una investigación? *Revista Uniandes Episteme*, 12(1), 1-3. *Revista Uniandes Episteme*, 12(1), 1-3. <https://dialnet.unirioja.es/descarga/articulo/10111456.pdf>
- Cornelio, C., Dash, S., Austel, V., Josephson, T. R., Goncalves, J., Clarkson, K. L., . . . Horesh, L. (2023). Combining data and theory for derivable scientific discovery with AI-Descartes. *Nature Communications*, 14(1), 1777. <https://doi.org/10.1038/s41467-023-37236-y>
- Cory, W. R., Cornelio, C., Dash, S., El Khadir, B., y Horesh, L. (2024). Evolving scientific discovery by unifying data and background knowledge with AI Hilbert. *Nature Communications*, 15(5922). <https://doi.org/10.1038/s41467-024-50074-w>
- Díaz, S. 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>
- Figuerola, P. (2025). Inteligencia artificial en la investigación académica y científica: ¿reemplazo, alianza o amenaza? *El Faro. Revista de Docencia Universitaria*, 2(2), 59-73. <https://doi.org/10.63790/xfmz3z44>
- Gamella, J., Peters, J., y Bühlmann, P. (2025). Causal chambers as a real-world physical testbed for AI methodology. *Nature Machine Intelligence*, 7. <https://doi.org/10.1038/s42256-024-00964-x>
- Ghareeb, A., Chang, B., Mitchener, L., Yiu, A., Szostkiewicz, C., Shved, D., . . . Rodriques, S. (2026). A multi-agent system for automating scientific discovery. *Nature*(655), 497–505. <https://doi.org/10.1038/s41586-026-10652-y>
- Gimeno, B. V., y Trigo, V. C. (2024). El rol de la inteligencia artificial en la publicación científica: perspectivas desde la farmacia hospitalaria. *Farmacia Hospitalaria*, 48(5), 246-251. <https://doi.org/10.1016/j.farma.2024.06.002>
- Girdwood, T., y Manikonda, L. (2026). Navigating artificial intelligence in qualitative research: ethical and practical considerations from concept to publication. *Frontiers in Research Metrics and Analytics*, 11, 1863790. <https://doi.org/10.3389/frma.2026.1863790>

- Gottweis, J., Weng, W.-H., Daryin, A., Tu, T., Sirkovic, P., Myaskovsky, A.,
Natarajan, V. (2026). Accelerating scientific discovery with Co-Scientist. *Nature*,
655, 487–496. <https://doi.org/10.1038/s41586-026-10644-y>
- Goyanes, M., y Lopezosa, C. (2024). ChatGPT en Ciencias Sociales: revisión de la
literatura sobre el uso de inteligencia artificial de OpenAI en investigación
cualitativa y cuantitativa. *Anuario ThinkEPI*, 18, e18e04.
<https://doi.org/10.3145/thinkepi.2024.e18a04>
- Guerrero, S. A., Ruiz, M. G., Yépez, G. D., y Sánchez, L. M. (2025). El impacto de la
inteligencia artificial en la producción científica. *Multidisciplinary Latin
American Journal*, 3(1), 629-649. <https://doi.org/10.62131/MLAJ-V3-N1-031>
- Henríquez, R. P., López, G. G., Domínguez, C. J., y Palazuelos, R. I. (2025). *Cultura
digital en contextos educativos: experiencias y significados*. Astra.
<https://doi.org/10.61728/AE20252175>
- Incio, F. F., Capuñay, S. D., Estela, U. R., Valles, C. M., Vergara, M. S., y Elera, G. D.
(2022). Inteligencia artificial en educación: una revisión de la literatura en revistas
científicas internacionales. *Apuntes Universitarios*, 12(1), 353-372.
<https://doi.org/10.17162/au.v12i1.974>
- Krenn, M., Pollice, R., Guo, S. Y., Aldeghi, M., Cervera-Lierta, A., Friederich, P., . . .
Aspuru-Guzik, A. (2022). On scientific understanding with artificial intelligence.
Nature Reviews Physics, 4, 761-769. <https://doi.org/10.1038/s42254-022-00518-3>
- Labañino, P. L., Lorca, M. A., De las Heras, P. M., y Campina, L. A. (2025). Evolución
del concepto de inteligencia artificial en la literatura científica: un análisis
sistemático. *Digital Education Review*, 46, 65-76.
<https://doi.org/10.1344/der.2025.46.65-76>
- Lalaleo, A. F., Carrera, C. F., y Martínez, Y. A. (2024). La IA como herramienta de apoyo
en la investigación científica en los docentes investigadores del ISTE. *Espíritu
Emprendedor TES*, 8(1), 97-110.
<https://www.espirituemprendedortes.com/index.php/revista/article/view/377/487>

- Li, M., Sun, J., y Tan, X. (2024). Evaluating the effectiveness of large language models in abstract screening: a comparative analysis. *Systematic Reviews*, 13, 219. <https://doi.org/10.1186/s13643-024-02609-x>
- Lopezosa, C. (2023). ChatGPT y comunicación científica: hacia un uso de la inteligencia artificial que sea tan útil como responsable. *Hipertext.net*, 26, 17-21. <https://doi.org/10.31009/hipertext.net.2023.i26.03>
- Lu, C., Lu, C., Lange, R., Yamada, Y., Hu, S., Foerster, J., . . . Clune, J. (2026). Towards end-to-end automation of AI research. *Nature*, 651, 914-919. <https://doi.org/10.1038/s41586-026-10265-5>
- Masbernat, P. (2024). Algunas reflexiones desde la revisión de literatura acerca de la ética en la investigación científica —integridad científica— en el contexto de la inteligencia artificial. *Instituto de Investigación en Derecho, Universidad Autónoma de Chile. SSRN Working Paper*. <https://ssrn.com/abstract=4707169>
- Messeri, L., y Crockett, M. (2024). Artificial intelligence and illusions of understanding in scientific research. *Nature*, 627, 49-58. <https://doi.org/10.1038/s41586-024-07146-0>
- Molina, I. L. (2024). Aspectos determinantes de la inteligencia artificial en la investigación educativa. *Praxis*, 20(3), 602-620. <https://doi.org/10.21676/23897856.6005>
- Montomoli, M. (2023). Un cambio de era: cómo la inteligencia artificial está revolucionando la publicación científica. *NefroPlus*, 15(1), 112-114. <https://www.revistanefrologia.com/es-download-pdf-X1888970023036820?local=true&log=false>
- Morantes, C. I. (2023). Inteligencia artificial en la investigación científica: sistematización y reflexiones sobre experiencias educativas. *Revista EDUCARE-UPEL-IPB-Segunda Nueva Etapa*, 27(3), 112-159. <https://doi.org/10.46498/reduipb.v27i3.2050>
- Muñoz, G. A., Villón, D. A., Calvachi, G. V., Orrala, Y. N., y Peñafiel, C. R. (2025). El impacto de la inteligencia artificial en la producción científica. *Ciencia y Reflexión. Revista Científica Multidisciplinaria*, 4(1), 632-667. <https://doi.org/10.70747/cr.v4i1.126>

- Nguyen, T. K. (2025). ChatGPT in thematic analysis: Can AI become a research assistant in qualitative research? *Quality & Quantity*, 59(6), 4945-4978. <https://doi.org/10.1007/s11135-025-02165-z>
- Ocampo, E. D., Vélez, J. D., y Gutiérrez-De Gracia, N. E. (2024). Tecnologías convergentes, impacto de la inteligencia artificial y las neurociencias en la formación de investigadores: una revisión sistemática. *Sociedad & Tecnología*, 7(S1), 210-231. <https://doi.org/10.51247/st.v7iS1.502>
- Perkins, M., y Roe, J. (2024). The use of generative AI in qualitative analysis: Inductive thematic analysis with ChatGPT. *Journal of Applied Learning and Teaching*, 7(1). <https://search.informit.org/doi/abs/10.3316/informit.T2025091900008591789301186>
- Rodríguez, F. E., Garcés, G. L., Valencia, J., y Valencia, A. A. (2025). Tendencias investigativas en el uso de técnicas de inteligencia artificial en la investigación científica. *Revista Venezolana de Gerencia*, 30(109), 351-380. <https://doi.org/10.52080/rvgluz.30.109.23>
- Román, A. D. (2023). Aplicación de la inteligencia artificial en la investigación académica: caso ChatGPT. *Revista Finanzas y Negocios*, 3(2), 41-61. <https://revistas.ulatina.edu.pa/index.php/Finanzasynegocios/article/view/323>
- Ruiz, M. G. (2024). Implicaciones de la inteligencia artificial en la metodología de investigación. *RITI Journal*, 12(26), 28-41. <https://doi.org/10.36825/RITI.12.26.003>
- Schwartz, M. (2022). Should artificial intelligence be interpretable to humans? *Nature Reviews Physics*, 4, 741-742. <https://doi.org/10.1038/s42254-022-00538-z>
- Shao, E., Wang, Y., Qian, Y., Pan, Z., Liu, H., y Wang, D. (2025). SciSciGPT: Advancing human–AI collaboration in the science of science. *Nature Computational Science*. <https://doi.org/10.1038/s43588-025-00906-6>
- Suazo, G. I. (2023). Inteligencia artificial en investigación científica. *SciComm Report*, 1(1), 1-3. <https://doi.org/10.32457/scr.v3i1.2149>

- Syriani, E., David, I., y Kumar, G. (2024). Screening articles for systematic reviews with ChatGPT. *Journal of Computer Languages*, 80, 101287. <https://doi.org/10.1016/j.cola.2024.101287>
- Szymanski, N., Rendy, B., Fei, Y., Kumar, R., He, T., Milsted, D., . . . Ceder, G. (2023). An autonomous laboratory for the accelerated synthesis of inorganic materials. *Nature*, 624, 86-91. <https://doi.org/10.1038/s41586-023-06734-w>
- Tang, X., Jin, Q., Zhu, K., Yuan, T., Zhang, Y., Zhou, W., . . . Gerstein, M. (2025). Risks of AI scientists: Prioritizing safeguarding over autonomy. *Nature Communications*. <https://doi.org/10.1038/s41467-025-63913-1>
- Tom, G., Schmid, S., Baird, S., Cao, Y., Darvish, K., Hao, H., . . . Aspuru-Guzik, A. (2024). Self-driving laboratories for chemistry and materials science. *Chemical Reviews*, 124(16), 9633-9732. <https://doi.org/10.1021/acs.chemrev.4c00055>
- Truhn, D., Eckardt, J.-N., Ferber, D., y Kather, J. N. (2024). Large language models and multimodal foundation models for precision oncology. *npj Precision Oncology*, 8(72). <https://doi.org/10.1038/s41698-024-00573-2>
- Van, K. A., Mahamune, U., Jongejan, A., Schaik, B. v., Balashova, D., Lashgari, D., . . . Moerland, P. (2024). ENCORE: A practical implementation to improve reproducibility and transparency of computational research. *Nature Communications*, 15(8117). <https://doi.org/10.1038/s41467-024-52446-8>
- Vimos, B. K., Viteri, O. J., Naranjo, S. M., y Novillo, H. K. (2024). Uso de la inteligencia artificial en los procesos de investigación científica, por parte de los docentes universitarios. *Journal of Economic and Social Science Research*, 4(4), 215-236. <https://doi.org/10.55813/gaea/jessr/v4/n4/143>
- Xie, Y., Sattari, K., Zhang, C., y Lin, J. (2023). Toward autonomous laboratories: Convergence of artificial intelligence and experimental automation. *Progress in Materials Science*, 132, 101043. <https://doi.org/10.1016/j.pmatsci.2022.101043>
- Zhang, Y., Khan, S., Mahmud, A., Yang, H., Lavin, A., Levin, M., . . . Zenil, H. (2025). Exploring the role of large language models in the scientific method: from hypothesis to discovery. *npj Artificial Intelligence*, 1(14). <https://doi.org/10.1038/s44387-025-00019-5>

INVESTIGACIÓN CIENTIFICA en la era de la INTELIGENCIA ARTIFICIAL



ISBN: 978-9907-848-06-9




CIDPROS
EDITORIAL