

De la IA como herramienta al IA Agéntica: cómo los agentes autónomos están redefiniendo la investigación científica

From AI as a tool to Agentic AI:
How autonomous agents are redefining scientific research

Gerardo Armando Picón 
Director Editor

RESUMEN

La inteligencia artificial ha transitado de ser una herramienta de apoyo a convertirse en sistemas agénticos autónomos capaces de planificar, ejecutar e iterar ciclos completos de investigación científica. Esta editorial analiza cómo dicha transición está transformando la producción de conocimiento y replantea la formación de investigadores en la educación superior. Se examinan sus aplicaciones más prometedoras, los riesgos de sesgo, autoría y desigualdad que conlleva, y se argumenta la necesidad de formar investigadores híbridos que articulen competencias humanas irremplazables con la orquestación responsable de agentes autónomos.

Palabras clave: IA Agéntica, Investigación Autónoma, Formación Científica



Este es un artículo publicado en acceso abierto bajo una Licencia Creative Commons.

ABSTRACT

Artificial intelligence has evolved from assistive tools to autonomous agentic systems capable of planning, executing, and iterating full scientific research workflows. This editorial examines how agentic AI is reshaping knowledge production, particularly in researcher training and higher education. It discusses promising applications, while addressing critical challenges related to research integrity, ethics, authorship, and access inequality. The text advocates for the development of hybrid researchers who combine irreplaceable human skills with the intelligent orchestration of autonomous agents. The future of science lies in responsible human-AI collaboration. (92 palabras)

Keywords: Agentic AI, Autonomous Research, Scientific Training

INTRODUCCIÓN

La inteligencia artificial (IA) ha irrumpido en la educación superior y la formación de investigadores con una velocidad sin precedentes. Sin embargo, el escenario más reciente va mucho más allá de las herramientas de apoyo al aprendizaje o la redacción académica: estamos ante el surgimiento de sistemas agénticos autónomos capaces de planificar y ejecutar flujos completos de investigación. Esta editorial examina qué significa la IA agéntica para la ciencia, qué oportunidades abre para la formación de investigadores y qué desafíos éticos, de integridad y equidad impone. Se concluye que avanzar responsablemente exige formar investigadores híbridos que mantengan el pensamiento crítico y el compromiso ético como ejes irrenunciables frente a la creciente autonomía de estos sistemas.

En nuestra última edición exploramos cómo la inteligencia artificial se consolida como una aliada estratégica en la formación de investigadores, ofreciendo herramientas para la revisión bibliográfica, el análisis de datos y la redacción académica. Ese panorama ha evolucionado de forma radical. Ya no hablamos solo de herramientas reactivas, sino de sistemas proactivos, autónomos y multiagente que pueden planificar, ejecutar e iterar flujos completos de investigación con mínima supervisión humana.

Este cambio representa uno de los saltos más significativos en la historia reciente de la ciencia. Investigaciones recientes coinciden en que los sistemas de IA agéntica permiten replicar en entornos in silico el trabajo colaborativo de equipos científicos completos, coordinando tareas especializadas con notable eficiencia (Collaco et al., 2026; Gridach et al., 2025; Schneider, 2025).

¿Qué es exactamente la IA Agénti-

ca?

A diferencia de los modelos generativos tradicionales, como ChatGPT o similares, que responden a prompts específicos, los sistemas agénticos poseen autonomía, razonamiento iterativo y capacidad de acción. Pueden descomponer un objetivo general, planificar pasos, buscar información, diseñar experimentos, analizar resultados simulados, ajustar hipótesis y producir un manuscrito preliminar.

Andrew Ng, uno de los principales referentes en el campo, ha enfatizado que los flujos de trabajo agénticos representan el verdadero avance productivo más allá de los grandes modelos de lenguaje. En su curso *Agentic AI (DeepLearning.AI)*, Ng describe patrones clave como la autocritica, el uso de herramientas externas y equipos de agentes especializados.

Wei et al. (2025) documentan la transformación de la ciencia hacia sistemas autónomos capaces de generar hipótesis, diseñar y ejecutar experimentos, e interpretar resultados. Su marco conceptual traza la evolución de la IA aplicada a la ciencia, identifica las capacidades clave de los sistemas agénticos y modela los flujos de trabajo del descubrimiento científico automatizado.

Aplicaciones transformadoras en la investigación actual

Los sistemas de IA con capacidad agéntica actúan como equipos computacionales autónomos que igualan el desempeño humano en tareas clave de investigación biomédica, como revisión bibliográfica, formulación de hipótesis y análisis de datos. Estos sistemas aceleran el descubrimiento de fármacos y biomarcadores mediante decisiones contextuales y retroalimentación experta, aunque persisten desafíos para su implementación generalizada (Li et al. 2026).

Sistemas como el “AI co scientist” de Google DeepMind y el “AI Scientist”

de Sakana AI combinan modelos de lenguaje de última generación con módulos especializados en diseño experimental, análisis de datos y búsqueda bibliográfica, lo que les permite identificar lagunas en el conocimiento, proponer hipótesis y protocolos experimentales detallados, y generar borradores coherentes de manuscritos; aunque aceleran ciclos de investigación y reducen tareas repetitivas, requieren supervisión experta para validar resultados, corregir sesgos y garantizar cumplimiento ético y regulatorio.

Oportunidades para la educación superior y la formación de investigadores

Este nuevo paradigma obliga a replantear la formación de investigadores. Si un agente agéntico puede ejecutar gran parte del trabajo rutinario y analítico, ¿qué competencias deben desarrollar los nuevos investigadores?

1. Pensamiento estratégico y validación crítica: La capacidad de formular preguntas científicas profundas y evaluar críticamente los outputs de los agentes.

2. Gobernanza ética y responsabilidad: Diseñar marcos de supervisión humana sobre sistemas autónomos.

3. Habilidades de orquestación: Aprender a construir y coordinar equipos de agentes personalizados.

4. Creatividad y juicio científico: Aquello que, hasta ahora, parece resistirse a la automatización completa.

Bengio et al. (2025) sostienen que los agentes de IA generalista con capacidad de planificación y acción autónoma conllevan riesgos catastróficos, por lo que proponen como alternativa más segura una "IA Científica" no-agencial. Este sistema, diseñado para explicar el mundo a partir de observaciones en lugar de actuar sobre él, combina un modelo del mundo que genera teorías y un motor de inferencia con manejo explícito de la incertidumbre. Su objetivo es asistir a investigadores humanos para

acelerar el progreso científico, incluyendo en la seguridad de IA, y servir como barrera de protección, favoreciendo así una ruta innovadora pero segura frente al desarrollo descontrolado de agentes.

Desafíos y riesgos que no podemos ignorar

El avance de la IA agéntica no está libre de riesgos significativos que la comunidad científica debe afrontar con rigor:

Según Hosseini et al. (2026) señalan que el uso creciente de agentes de IA en tareas científicas mejora la productividad, pero también genera múltiples riesgos: desde decisiones políticas basadas en resultados erróneos o sesgados, hasta la pérdida de empleos y habilidades humanas, pasando por lagunas de responsabilidad y participación en investigaciones poco éticas. Para una gestión responsable, es urgente definir qué tareas no deberían automatizarse. Las instituciones deben capacitar a los investigadores en alfabetización algorítmica y designar roles específicos, como un garante de IA, que supervisen y aseguren la integridad de las contribuciones asistidas por IA.

Las IA agénticas pueden amplificar sesgos y generar conclusiones falsas pero convincentes, un fenómeno conocido como alucinación algorítmica. Este tipo de IA multiplica los riesgos de sesgo al perpetuar discriminaciones históricas presentes en sus datos de entrenamiento, amplificar sesgos cognitivos humanos y crear una dependencia peligrosa que reduce el rigor y la revisión crítica por parte del investigador. Además, la autonomía de la IA agéntica significa que puede solicitar acceso a sistemas y datos externos; a medida que los agentes procesan grandes volúmenes de información con velocidad, aumenta la posibilidad de encontrar información confidencial. Abordar la privacidad y el cumplimiento normativo, así como garantizar información accesible sobre las acciones y

los datos a los que se accede, pueden permitir que la función de IA agéntica asuma menos riesgos en relación con la información confidencial (Chen, 2025).

La autonomía de la IA agéntica plantea interrogantes fundamentales sobre la autoría: ¿quién es el autor legítimo de un artículo científico generado mayoritariamente por agentes de inteligencia artificial? ¿Se considera al investigador que orquestó el proceso, a los desarrolladores del sistema, o al propio agente? Esta cuestión es particularmente compleja porque los agentes pueden realizar tareas que tradicionalmente corresponden al autor humano, desde la búsqueda bibliográfica hasta el análisis de datos y la redacción de secciones completas. Las revistas científicas están comenzando a actualizar sus políticas de publicación para abordar esta realidad emergente, pero aún queda mucho terreno por definir. La falta de claridad normativa genera incertidumbre sobre los derechos de propiedad intelectual, el reconocimiento académico y la atribución de responsabilidad en caso de errores o irregularidades en el contenido publicado.

La adopción de IA agéntica en la investigación científica amenaza con profundizar las desigualdades existentes entre instituciones. Las universidades y centros de investigación con mayor acceso a estos sistemas avanzados, generalmente aquellos con recursos financieros sustanciales, infraestructura tecnológica robusta y personal capacitado, podrían generar una brecha aún mayor respecto a aquellas con recursos limitados. Esta disparidad podría traducirse en ventajas competitivas significativas en términos de productividad investigadora, calidad de las publicaciones, capacidad de análisis de datos complejos y velocidad de descubrimiento científico. En consecuencia, se corre el riesgo de consolidar un escenario donde la excelencia científica quede reservada para una élite insti-

tucional, mientras que las investigaciones de instituciones con menos recursos queden en desventaja estructural, perpetuando y exacerbando las asimetrías históricas en la producción de conocimiento a nivel global.

No se trata de frenar el avance ni de seguirlo sin cuestionarlo, sino de conducirlo con criterio y responsabilidad. Las universidades deberían:

- Modernizar los programas de postgrado incorporando formación específica sobre agentes de IA.
- Establecer normativas institucionales que regulen el uso ético de estos sistemas.
- Impulsar la colaboración interdisciplinaria que integre ciencias, ética, derecho y filosofía.
- Exigir trazabilidad y transparencia garantizando que cada resultado generado por un agente incluya un registro de procedencia detallado.

La ciencia del mañana no enfrentará a personas con algoritmos, sino que los integrará, siempre que sostengamos firmes el timón en principios irrenunciables como la verdad, la integridad y el bien común.

Un nuevo contrato social con la IA

Estamos presenciando el nacimiento de una nueva forma de hacer ciencia. La IA agéntica promete acelerar el descubrimiento como nunca antes, pero exige a cambio una madurez ética y una reinención de la formación científica que aún estamos empezando a imaginar.

No se trata solo de adoptar una nueva tecnología, sino de redefinir qué significa ser científico en el siglo XXI. El futuro ya no es ciencia hecha por humanos con ayuda de máquinas. Es ciencia hecha por equipos híbridos donde la agencia se distribuye de forma inteligente. La pregunta que debemos responder como comunidad académica es: ¿estamos preparados para liderar esta transformación?

REFERENCIAS

- Bengio, Y., Cohen, M., Fornasiere, D., Ghosh, J., Greiner, P., MacDermott, M., Mindermann, S., Oberman, A., Richardson, J., Richardson, O., Rondeau, M.-A., St-Charles, P.-L., & Williams-King, D. (2025). Superintelligent Agents Pose Catastrophic Risks: Can Scientist AI Offer a Safer Path? (arXiv:2502.15657). arXiv. <https://arxiv.org/abs/2502.15657>
- Chen, M. (2025, 17 de junio). ¿Qué es la IA agéntica? Oracle América Latina. <https://www.oracle.com/latam/artificial-intelligence/agent-ai/>
- Collaco, B.G., Haider, S.A., Prabha, S., Gomez-Cabello, C. A., Genovese, A., Wood, N., Bagaria, S. P., Gopala, N., Tao, C., & Jorge Forte, F. (2026). The role of agentic artificial intelligence in healthcare: a scoping review. *npj Digit. Med.* 9, 345. <https://doi.org/10.1038/s41746-026-02517-5>
- Hosseini, M., Murad, M., & Resnik, D. B. (2026). Benefits and Risks of Using AI Agents in Research. *The Hastings Center report*, 56(1), 13–17. <https://doi.org/10.1002/hast.70025>
- Gridach, M., Nanavati, J., Zine El Abidine, K., Mendes, L., & Mack, C. (2025). Agentic AI for scientific discovery: a survey of progress, challenges, and future directions. arXiv preprint <https://doi.org/10.48550/arXiv.2503.08979>
- Li, B., Saini, A.K., Hernandez, J.G. & Moore, J.H. (2026). Agentic AI and the rise of in silico team science in biomedical research. *Nat Biotechnol*, 44, 711–725. <https://doi.org/10.1038/s41587-026-03035-1>
- Ng, A. (2025-2026). Agentic AI, DeepLearning.AI. <https://www.deeplearning.ai/courses/agent-ai>
- Schneider, J. (2025). Generative to Agentic AI: Survey, Conceptualization, and Challenges. <https://doi.org/10.48550/arXiv.2504.18875>
- Wei, J., Yang, Y., Zhang, X., Chen, Y., Zhuang, X., Gao, Z., Zhou, D., Wang, G., Gao, Z., Cao, J., Qiu, Z., Hu, M., Ma, C., Tang, S., He, J., Song, C., He, X., Zhang, Q., You, C., ... Zhou, B. (2025). From AI for Science to Agentic Science: A Survey on Autonomous Scientific Discovery (arXiv:2508.14111). arXiv. <https://arxiv.org/abs/2508.14111>