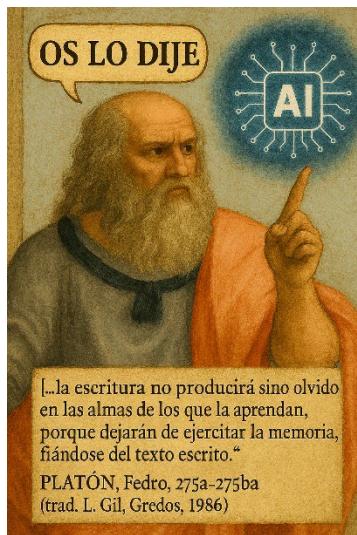


La autoría humana en la era de la inteligencia artificial

Manifiesto de pertinencia y uso ético de las IAs en la producción académica en la rama de las humanidades

Dr. Eduardo Zamarro, diciembre 2025

Orcid ID: [0000-0001-5512-2652](https://orcid.org/0000-0001-5512-2652)



ChatGPT. (2025). *Imagen generada con IA basada en la obra “La Escuela de Atenas” de Rafael, con texto adaptado de Platón, Fedro, 275a-275b (trad. L. Gil, Gredos, 1986)* [Imagen]. OpenAI.

Un manifiesto filosófico y deontológico sobre la imposibilidad de la IA como autora y la defensa de la autonomía del sujeto humano

La tesis central de este manifiesto es que solo los seres humanos pueden ser autores, porque solo ellos son autónomos, pueden darse leyes a sí mismos, exponerse en la verdad de una obra y responder por lo que dicen ante otros. La inteligencia artificial, por el contrario, permanece ligada a *prompts*, datos y sesgos ajenos, y produce “no-cosas” informacionales sin mundo propio; por eso no puede ser autora, sino únicamente herramienta sometida al juicio humano.

1. Fundamentación filosófica de la autoría humana

En la *Fundamentación de la metafísica de las costumbres*, Kant afirma que la autonomía de la voluntad es fundamento de la dignidad de la naturaleza humana y de toda naturaleza racional, lo que implica que solo un ser capaz de auto-legislarse moralmente puede ser sujeto de deberes y responsabilidades (Kant, 2005/1785). Como recuerda Kant, la autonomía, entendida como capacidad de “darse ley a sí mismo”, es la raíz de la dignidad humana, y por ello también de cualquier pretensión legítima de autoría. Una IA, que actúa siempre bajo reglas y fines impuestos desde fuera, permanece en la heteronomía y no puede, por tanto, ocupar el lugar del autor responsable en sentido fuerte; la autoría, en este manifiesto, se reserva estrictamente a sujetos humanos.

En *El origen de la obra de arte*, Heidegger concibe el arte como el lugar en que la verdad acontece como desocultamiento (*aletheia*): la obra abre un mundo y deja comparecer a los entes en su ser (Heidegger, 2019/1935–1936). Esta concepción exige un “quien” que se expone en la obra, que se arriesga en la puesta en obra de una verdad; una máquina puede contribuir

técnicamente a la producción de formas, pero no puede asumir la apertura de un mundo que compromete un modo humano de estar-en-el-mundo.

Platón, en la alegoría de la caverna (*República*, libro VII, 514a-517c), describe la diferencia entre quienes toman sombras por realidad y quien emprende el giro hacia la luz del Bien, imagen del conocimiento auténtico (Platón, 1986/ca. 380 a. C.). En el *Fedro* (274b–278b), mediante el mito de Theuth y Thamus, critica la escritura porque, al ofrecer signos externos, favorece una ilusión de saber y debilita la memoria viva: quien se apoya en la escritura puede “parecer” sabio sin serlo realmente (Platón, 2016/ca. 370 a. C.). En este manifiesto se asume que la IA se parece más a esta escritura sospechosa que a la memoria viva: produce huellas sin experiencia propia, sin conciencia de verdad, sin mundo.

En *No-cosas. Quiebras del mundo de hoy*, Byung-Chul Han describe la transición desde un mundo de cosas —accesibles al cuerpo, resistentes, habitables— a un entorno dominado por flujos de información digitales, desmaterializadas, que llenan el mundo de datos inquietantes (Han, 2021). La IA generativa aparece entonces como un dispositivo paradigmático de producción de este tipo de información sin cuerpo, que requiere más que nunca un sujeto humano que la sitúe en un mundo de sentido.

2. Manifiesto sobre la autoría humana y el uso de IA

1. El sujeto humano siempre es el autor

La autoría implica responder por lo que se dice y hace; solo un ser capaz de darse a sí mismo una ley moral y de reconocer el deber como propio puede ser autor en sentido fuerte (Kant, 2005/1785). La IA, que carece de libertad práctica, de conciencia de deber y de experiencia vital, no será considerada autora ni coautora en ningún trabajo que se suscriba bajo este manifiesto: únicamente los sujetos humanos pueden asumir la responsabilidad última del contenido de una obra o investigación.

2. El arte como desvelamiento exige un “quién”, no solo un “qué”

Si el arte es, con Heidegger, un modo privilegiado en que la verdad acontece como desocultamiento, toda obra requiere un agente que se juegue en ese abrir un mundo (Heidegger, 2019/1935–1936). La IA puede intervenir técnicamente en la elaboración material de una obra, pero no puede asumir la exposición ontológica de poner en obra una verdad de la existencia; la autoría artística y hermenéutica permanece, por tanto, como tarea humana, especialmente en el ámbito de las humanidades.

3. La IA en la caverna de los datos y la memoria viva

Como los prisioneros de la caverna platónica, la IA solo manipula sombras: correlaciones extraídas de un inmenso muro de datos (Platón, 1986/ca. 380 a. C.). Solo un sujeto humano puede emprender el giro hacia la luz, reconocer la diferencia entre apariencia y realidad y hacerse responsable del camino educativo que implica salir de la caverna. En el *Fedro*, Platón sostiene que la escritura produce olvido en las almas, porque, fiándose de lo escrito, los hombres descuidarán la memoria y el saber vivo que puede ser defendido ante otros (Platón, 2016/ca. 370 a. C.). La IA, como hiper-escritura automatizada, corre el riesgo de intensificar ese olvido de sí, sustituyendo la experiencia biográfica y narrativa por una acumulación de textos que nadie asume como propios; este manifiesto rechaza esa delegación de la memoria, del pensamiento y del aprendizaje en sistemas automáticos. ¿Qué sería de los avances en todas las ciencias si hubiéramos denostado la escritura por ser una tecnología perversa?

4. Responsabilidad, fines y experiencia: la IA como herramienta subordinada

Ninguna decisión, afirmación o juicio generado con ayuda de IA queda descargado en el sistema: la responsabilidad recae siempre en quienes lo diseñan, lo entrenan y lo

utilizan (UNESCO, 2024/2021). La IA no decide fines, solo optimiza medios; la libertad práctica consiste en poder elegir fines a la luz de razones, mientras que el sistema se limita a ejecutar objetivos fijados desde fuera. Sin capacidad de cuestionar esos fines ni de reformular “valores” propios, y sin nacimiento, muerte, sufrimiento, goce o memoria vivida, la IA no puede ser portadora de un “punto de vista” en sentido fuerte. Por ello, todo uso de IA en escritura, investigación, arte o docencia ha de ser transparente y subordinado a un criterio humano explícito: la máquina solo puede ser herramienta declarada, nunca instancia decisoria ni sujeto de autoría.

5. **Autonomía, investigación en humanidades y exigencia de rigor**

En el ámbito de las humanidades —filosofía, arte, historia, filología, estudios culturales— el núcleo del trabajo de investigación reside en la lectura situada, la interpretación crítica, la articulación conceptual y la construcción de una voz propia. Usar IA de forma acrítica para producir textos “correctos” pero vacíos supone renunciar a esa tarea y empobrecer tanto el campo como la formación de quienes investigan. En línea con la idea kantiana de madurez ilustrada como salida de la minoría de edad, este manifiesto defiende que la IA solo es éticamente aceptable en las humanidades cuando ayuda a ampliar las capacidades de comprensión, argumentación y creación del sujeto humano, nunca cuando las sustituye o las debilita.

6. **Crítica fácil, “cuñaos” y “haters”: contra la falsa superioridad moral**

El debate actual sobre la IA en las humanidades se ve a menudo colonizado por la figura del “cuñao digital ilustrado” y del “hater profesional”: personajes que, tras ver dos vídeos en redes y leer ningún libro, se sienten en condiciones de sentenciar el fin de la cultura, de la universidad y, si se tercia, de la especie humana. La crítica de la IA se convierte así en la vía rápida para obtener autoridad moral a bajo coste: basta con repetir un par de tópicos (“la máquina nos quitará el trabajo”, “lo auténtico es lo analógico”) para ganar aplausos de barra de bar sin necesidad de hacerse cargo de la complejidad técnica, política y filosófica del problema.

Este manifiesto rechaza tanto la demonización simplista como la celebración ingenua de la IA. Frente al “cuñadismo” y la cultura del odio, defiende una crítica informada y responsable: leer los textos, conocer mínimamente el funcionamiento de las herramientas, situar el debate en la historia de las tecnologías de la escritura y del conocimiento, y evaluar caso por caso la pertinencia del uso de IA. La verdadera postura crítica no consiste en descalificar desde fuera, sino en asumir la tarea —mucho menos vistosa, pero infinitamente más seria— de pensar con rigor qué lugar puede o no puede tener la IA en la investigación en humanidades, preservando siempre la dignidad de la autoría humana.

3. Respaldo institucional

Distintas instituciones académicas de prestigio respaldan esta tesis al negar a la IA la condición de autora y mantener la responsabilidad en manos humanas. El *Ombudsman* de la investigación en Alemania, instancia colegiada que da referencia a universidades e institutos, afirma que la IA no puede ser autora porque no puede asumir responsabilidad por un manuscrito ni aprobar su versión final (Ombudsman für die Wissenschaft, 2023). Este criterio se usa como guía para políticas de integridad científica en universidades alemanas, reforzando la idea de que la autoría exige capacidad de responder y de consentir.

La Universidad de Glasgow señala en su guía para investigadores que las herramientas de IA “no saben lo que producen” y que cualquier uso de IA debe ir acompañado de análisis crítico y supervisión por parte de quien investiga (University of Glasgow, 2025). Además, advierte que el

trabajo presentado como propio pero generado por IA sin reconocimiento no cumple los requisitos de integridad académica y de investigación, insistiendo en que el esfuerzo intelectual y la responsabilidad deben seguir siendo humanos.

Las recomendaciones del *International Committee of Medical Journal Editors* (ICMJE) —adoptadas por muchas facultades y hospitales universitarios— establecen que los sistemas de IA no deben figurar como autores ni coautores, porque no pueden hacerse responsables de la exactitud, integridad y originalidad del trabajo (ICMJE, 2024). Los autores humanos siguen siendo plenamente responsables de todo el contenido, incluso de las partes generadas con ayuda de IA, y deben declarar expresamente cualquier uso de estas herramientas. Estas políticas institucionales convergen con el núcleo del presente manifiesto: la IA puede ser recurso técnico, pero la autoría y la responsabilidad permanecen necesariamente en manos humanas.

Bibliografía

- Heidegger, M. (2019). *El origen de la obra de arte* (Trad. J. L. Vermal). LaOficina. (Trabajo original publicado 1935–1936)
- Han, B.-C. (2021). *No-cosas. Quiebras del mundo de hoy* (Trad. J. Chamorro Mielke). Taurus.
- ICMJE. (2024). *Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals*. (Documento original en línea, versión 2024)
- Kant, I. (2005). *Fundamentación de la metafísica de las costumbres* (Trad. C. García Morente & C. García Trevijano). Tecnos. (Trabajo original publicado 1785)
- Platón. (1986). *República* (Libros VI–VII) (Trad. J. M. Pabón & M. Fernández-Galiano). Alianza. (Trabajo original publicado ca. 380 a. C.)
- Platón. (2016). *Fedón. Fedro* (Trad. L. Gil). Alianza. (Trabajos originales publicados ca. 399–370 a. C.)
- UNESCO. (2024). *Recomendación sobre la Ética de la Inteligencia Artificial*. UNESCO. (Documento original aprobado 2021)
- University of Glasgow. (2025). *Generative AI guidance for researchers*. (Documento original en línea, versión 2025)

Declaración de uso de IA

En la elaboración de este trabajo se utilizaron herramientas de inteligencia artificial generativa como apoyo instrumental en la búsqueda preliminar de información y en la reformulación de borradores (Perplexity) y en la identificación de literatura académica relevante (Scopus AI). El contenido ha sido revisado críticamente por el autor y contrastado con las fuentes citadas, por lo que la responsabilidad última de los argumentos y conclusiones recae exclusivamente en el autor.



Eduardo Zamarro
Dr. en Bellas Artes