
TRIBUNA BiD

Uso ético y eficiente de la inteligencia artificial en trabajos académicos: Veritas e interacción crítica escalonada

1

Ethical and efficient use of artificial intelligence in academic work: Veritas and staggered critical interaction

Ús ètic i eficient de la intel·ligència artificial en treballs acadèmics: Veritas i interacció crítica escalonada

Lluís Codina 

© Autor

Universitat Pompeu Fabra

lluis.codina@upf.edu

Recibido: 03-12-2025

Aceptado: 03-12-2025

Citación recomendada

Codina, Lluís (2025). Uso ético y eficiente de la inteligencia artificial en trabajos académicos: Veritas e interacción crítica escalonada. *BiD*, 55. <https://doi.org/10.1344/BID2025.55.01>

Resumen

Presentación de los puntos cardinales del uso de la inteligencia artificial: las IA no pueden ser acreditadas como autoras; la responsabilidad de un trabajo académico es íntegra y exclusiva del autor; las personas autoras deben aplicar el pensamiento crítico en el uso de las IA y el uso significativo de la IA debe ser declarado. Consecuencias de estos puntos cardinales aplicados al uso de la IA en la producción de trabajos académicos evaluativos. Presentación del marco Veritas: verificación, evaluación, atribución y edición como procedimiento para transformar y filtrar contenidos generados por una IA antes de aplicarlo en un trabajo académico. Presentación de la interacción crítica escalonada como la filosofía detrás de Veritas para el uso ético de la IA pero también para mitigar o anular la deuda cognitiva.

Palabras clave

Inteligencia artificial generativa; trabajos académicos evaluativos; ética; Veritas; interacción crítica escalonada.

Abstract

Presentation of the cardinal points of artificial intelligence use: AIs cannot be credited as authors; the responsibility for an academic work rests entirely and exclusively with the author; authors must apply critical thinking in the use of AI, and the significative use of AI must be declared. Consequences of these cardinal points applied to the use of AI in the production of evaluative academic work. Presentation of the Veritas framework: verification, evaluation, attribution, and editing as a procedure for transforming and filtering AI-generated content before applying it in academic work. Presentation of staggered critical interaction as the philosophy behind Veritas for the ethical use of AI, but also for mitigating or eliminating cognitive debt.

Keywords

Generative artificial intelligence; evaluative academic work; ethics; Veritas; staggered critical interaction.

2

Resum

Presentació dels punts cardinals de l'ús de la intel·ligència artificial: les IA no poden ser acreditades com a autors; la responsabilitat d'un treball acadèmic és íntegra i exclusiva de l'autor; les persones autors han d'aplicar el pensament crític a l'ús de les IA i l'ús significatiu de la IA ha de ser declarat. Conseqüències d'aquests punts cardinals aplicats a l'ús de la IA en la producció de treballs acadèmics avaluatius. Presentació del marc Veritas: verificació, avaluació, atribució i edició com a procediment per transformar i filtrar continguts generats per una IA abans d'aplicar-ho a un treball acadèmic. Presentació de la interacció crítica esglaonada com la filosofia darrere de Veritas per a l'ús ètic de la IA però també per mitigar o anul·lar el deute cognitiu.

Paraules clau

Intel·ligència artificial generativa; treballs acadèmics avaluatius; ètica; Veritas; interacció crítica esglaonada.

1. Introducción

Como es bien sabido, en pocos años, la utilización de la inteligencia artificial generativa en trabajos académicos se ha vuelto casi universal o está camino de serlo. Antes de seguir, a efectos de lo que seguirá, debo aclarar que, por trabajos académicos entiendo los trabajos académicos que tienen un carácter **evaluativo**.

Para que se entienda enseguida, señalaré que, entre los más representativos de estas producciones encuentran los trabajos de fin de máster y las tesis doctorales. Podemos añadir, sin mayor problema (todo lo contrario), los trabajos finales de grado que también exhiben en muchas ocasiones una notable calidad.

Lo importante, sobre todo en el caso de las tesis doctorales, es que los trabajos académicos de mayor calidad realizan aportaciones muy relevantes con implicaciones tanto académicas como sociales.

Además, precisamente por el carácter evaluativo al que me he referido, cuando se superan, dan derecho a titulaciones que gozan de amplia consideración social. En concreto, el trabajo final de máster genera magisters y las tesis doctorales, doctores, lo que abre nuevas posibilidades profesionales a los poseedores de estos títulos.

Una característica adicional de esta clase de títulos académicos es que la sociedad espera de ellos que sean otorgados de forma justa, esto es, que los tribunales respectivos solamente los concedan a quienes realmente los han merecido con su esfuerzo y no, por ejemplo, con el uso de la inteligencia artificial sustituyendo el esfuerzo del autor.

Lo que voy a discutir aquí se refiere por tanto a esta clase de producciones académicas. No obstante, buena parte de los que se expondrá creo que es extrapolable a otros formatos de la comunicación académicas, como artículos científicos, dado que, aunque sea por otras vías, también constituyen de alguna forma trabajos evaluativos. Por ejemplo, en muchos países la publicación de artículos científicos conduce a mejoras en la contratación académica.

Por último, me gustaría explicar como se ha gestado esta Tribuna. Desde hace dos años, he impartido sesiones de formación predoctoral o seminarios para expertos en mi universidad (Pompeu Fabra) pero también en universidades de Cataluña, de España y de América Latina. La necesidad de una forma operativa de enseñar y de discutir el uso de la IA forjó las bases del marco de procedimientos que llamo Veritas y que presento aquí. Para llegar a este marco, durante estos dos años, me he ido documentado extensivamente usando tanto buscadores convencionales como sistemas de IA. Por último, las experiencias formativas y el debate con colegas en seminarios, me ha ido ayudando a refinar y presentar Veritas de un modo que espero sea a la vez tan coherente como práctico.

2. Los puntos cardinales de la ética de la IA aplicada a la academia

Los responsables de titulaciones universitarias, desde los grados a los programas de doctorados saben bien que, a día de hoy, existen numerosas guías y orientaciones sobre el uso de la inteligencia artificial en la academia, no solo en trabajos evaluativos, sino también (y sobre todo) en investigación y en publicación científica en general (ver al respecto, la sección de Referencias).

Y este es el punto al que quería llegar. A poco que el lector haya examinado algunas de estas guías o normativas, se dará cuenta de que, o bien tienen un carácter abstracto o bien cuando descienden al detalle, dejan zonas que requieren interpretación.

Afortunadamente, lo anterior no impide que emerja un claro núcleo duro de consenso de una lectura atenta de las mismas. Este consenso emergente se manifiesta en, al menos, los cuatro puntos cardinales que se van a presentar (recuerden que un punto cardinal no es necesariamente un punto geográfico). Así mismo, estos puntos solapan en general con las disposiciones éticas de las revistas científicas.

Los enuncio a continuación de una forma que espero que el lector interesado en estos temas pueda reconocer fácilmente:

- 1. No autoría.** Las IA no pueden ser acreditadas ni como autoras ni como coautoras. Esta no-autoría desaconseja citar a la IA como fuentes de autoridad. En su lugar, se promueve la cita de la fuente original.
- 2. Responsabilidad.** La responsabilidad del contenido de un trabajo académico corresponde íntegramente a la persona autora.
- 3. Pensamiento crítico.** Las personas autoras deben utilizar la IA aplicando su pensamiento crítico para la detección de “alucinaciones”, evitación de sesgos y prevención de errores o inconsistencias.
- 4. Transparencia.** El uso significativo de la IA en un trabajo académico, debe ser declarado tanto de forma transparente como trazable.

Son cuatro puntos que a veces se reducen a los dos primeros, por una buena razón. La combinación del primero y el segundo, bajo el marco general de las buenas prácticas científicas, imponen el tercero de forma natural, y el cuarto, por su parte, es una exigencia de propia de la tradición académica general, pues el uso de la IA no anula las buenas prácticas convencionales. Sin embargo, por claridad prefiero exponer los 4 puntos cardinales en su forma más explícita.

3. La oferta de sistemas de inteligencia artificial para la academia

Un dilema aparente surge, por las razones que se verán, cuando consideramos las características de los sistemas de inteligencia artificial orientados a su uso en el entorno académico. Me refiero a productos, algunos bien conocidos, como los siguientes cuatro que señalaré, a modo de definición ostensiva:

- Elicit
- SciSpace
- Consensus
- Undermind

Se trata de software de inteligencia artificial que se ha especializado claramente en producir lo que podemos llamar **síntesis de la evidencia** (marcos teóricos, revisiones de la literatura, etc.). A partir de aquí están orientados (no hay más que ver sus funciones) en generar contenidos con un formato final genuinamente académico, de tal forma que un usuario *podría* (véase el condicional) utilizar directamente como parte o la totalidad una o más secciones de su trabajo.

Pero, si volvemos la vista a los puntos cardinales enunciados, debemos dudar de esa posibilidad. La razón es que al aplicar estos puntos cardinales a los trabajos académicos, en concreto los puntos 1 y 2 sobre todo, pero también el 4, aparece una consecuencia lógica usando tanto la deducción como la reducción al absurdo. Esta consecuencia es la siguiente:

- Ninguna sección de un trabajo académico evaluativo puede ser escrita de modo directo, ni totalmente ni parcialmente, por una inteligencia artificial.

Veamos por qué: el punto 1, al prohibir añadir el nombre de una IA a la *by line*, por lógica, también nos prohíbe usar contenido de una IA como si fuera propio. La razón: si (a) ese contenido no es de los autores, y (b) tampoco se puede atribuir a la IA, entonces (c) no puede utilizarse como si fuera propio sin cometer plagio.

El punto 2 actúa como un refuerzo: dado que (a) si usamos una IA en alguna fase (admisible) de nuestro trabajo, y (b) ante cualquier error o incluso ante cualquier problema legal, solamente podemos ser responsables los autores (humanos), se deduce (c) que el uso directo, esto es, sin ninguna clase ni de supervisión ni de filtrado, queda lógicamente desacreditado.

El punto 4, por su parte, completa el círculo lógico al exigir el uso del pensamiento crítico para poder verificar, analizar y, en su caso filtrar, los contenidos generados por la IA. ¿Y qué sucede con el punto 3? Forma parte, por un lado de las exigencias que vienen de la mano del nuevo escenario de uso de la IA, pero al mismo tiempo, forma parte de las exigencias tradicionales de todo trabajo académico a la hora de aportar los elementos de transparencia y trazabilidad habituales.

4. El problema

Plateado todo lo anterior, y en forma de pregunta, aparece el siguiente problema: ¿de qué forma puede un investigador pre o postdoctoral poner a su servicio de forma legítima las ventajas de la IA a la hora de aplicarla a trabajos académicos, en especial si son de tipo evaluativos?

La forma de responder al problema es que, si un contenido generado por un IA se considera útil, debe ser **transformado** de forma que su aplicación se reconcilie con la ética esperada de los trabajos evaluativos y que he condensado en los 4 puntos cardinales. Y aquí aparece la propuesta de trabajo o de transformación que llamo Veritas.

5. Veritas

5

La función de Veritas es aportar un marco de trabajo que permita utilizar contenidos generados por una IA generativa en trabajos académicos que sea compatible con los 4 puntos cardinales de la ética de la IA en la academia. Veritas se compone de cuatro acciones que, si las miramos bien, en realidad son características del buen hacer académico habitual. Son las siguientes:

- Verificar
- Evaluar
- Atribuir
- Editar

El desarrollo de los elementos de Veritas es como sigue:

- **Verificar las fuentes.** Debe comprobarse cualquier fuente candidata a ser usada y citada, si se reputa de interés para el trabajo académico. Esta verificación incluye la totalidad de las referencias de interés, ya sea para apoyar ideas o alguna clase de datos que pensemos manejar en nuestro trabajo.
- **Evaluar los constructos** (ideas, conceptos, modelos, explicaciones, etc.) presentados por la IA, mediante pensamiento crítico. En esta evaluación, el autor (el doctorando, p.e.) no solo debe asegurarse de captar la esencia de los mismo, sino que debe ponerlos en relación con los conocimientos previos del investigador en su materia, así como con otros elementos de su marco teórico o conceptual.
- **Atribuir las ideas de terceros.** Cualquier idea o contenido valioso que el investigador consiga de las dos operaciones previas, deben ser atribuidos a las fuentes originales (a los autores humanos, no a la IA) en la medida que considere incorporarlos a su trabajo. Las atribuciones se harán o bien por explicación o parafraseo (más la cita parentética adecuada) o bien por cita literal (marcando y delimitando el fragmento literal más la cita parentética).
- **Editar el contenido para garantizar su coherencia.** Cualquier contenido producido con asistencia de una IA, salvo las citas literales, debe ser editado por el autor para generar un discurso coherente, unificar la escritura y dotar a toda la narrativa de su trabajo de una voz propia. Esto, además de dotar a la tesis de un estilo coherente, garantiza que el autor ha asumido su responsabilidad y está en condiciones de explicar todos los conceptos que ha manejado en su trabajo.

Ahora podemos ver en la siguiente tabla la relación entre los puntos cardinales y los elementos de Veritas (Tabla 1).

Tabla 1. Veritas y los puntos cardinales en el uso de la IA

Puntos cardinales	Puntos Veritas
Las IA no son autoras	Evaluación, Verificación
Las IA no pueden ser citadas como autoras	Atribución
El autor es plenamente responsable de los contenidos	Evaluación, Verificación, Edición
El autor debe usar la IA con pensamiento crítico	Evaluación, Verificación

Fuente: elaboración propia

6. Interacción crítica escalonada

La filosofía que sustenta Veritas la denomino **interacción crítica escalonada (ICE)**. Es equivalente al importante principio denominado HITL, por la expresión *human-in-the-loop*, pero creo que más expresiva. HITL es un concepto procedente de la ingeniería de la IA, y sostiene que la persona interviene activamente siempre en el ciclo de funcionamiento de un sistema de IA para supervisar, corregir y tomar las decisiones clave.

ICE implica que el autor no se limita a usar contenidos generados por una IA de forma pasiva, o “por copiar y pegar”. De hecho, es lo diametralmente opuesto a estas dos cosas. Es también la única forma segura de mitigar o incluso anular la **deuda cognitiva**. Recordemos que se genera deuda cognitiva cuando una persona humana delega funciones intelectuales o cognitivas avanzadas a una IA. El problema principal es que, de la deuda cognitiva pasamos a la **atrofia cognitiva**. No parece algo que pueda desear un investigador en formación (ni tampoco un senior).

La relación entre la ICE y Veritas es como sigue: la primera (ICE) es la filosofía general que debe presidir el uso de la IA en la academia (tal vez el uso de la IA en general); la segunda (Veritas) es una de las formas operativas en las cuales se puede concretar. Un desarrollo comprensivo del significado y las implicaciones de ICE es como sigue:

- **Interacción.** Implica que el investigador no toma los resultados de una IA de forma pasiva. Por el contrario, interactúa tanto con la IA como con sus resultados. Con la IA cuestionando sus resultados y repreguntando, redirigiendo sus respuestas, etc. Con los resultados usando el escrutinio.
- **Crítica.** La interacción anterior debe estar guiada por el pensamiento crítico, lo que obliga al investigador a aplicar, entre otras cosas, marcos de actuación como Veritas.
- **Escalonada.** Es escalonada porque procede por fases, en función de las necesidades de uso de la IA y de las acciones de verificación, etc., que se deben realizar, de las repreguntas, etc. También implica que el investigador está dispuesto a invertir esfuerzo intelectual en procesos de verificación, etc.

7. Declaración transparente y trazable

Nos queda el último de los puntos cardinales, el que señala: “4. Transparencia. El uso significativo de la IA en un trabajo académico, debe ser declarado tanto de forma transparente como trazable”.

Aquí corresponde señalar que este punto no procede del uso de la IA, sino de la buena práctica habitual de cualquier trabajo académico, más aún de trabajos evaluativos donde los autores deben declarar los métodos por los cuales han obtenido sus resultados. Pero esta declaración en sí misma, que responde al principio de la transparencia, es

necesaria pero no suficiente. Cuando usamos una metodología de forma significativa debemos aportar además datos de trazabilidad.

En este punto, lo que corresponde al autor es aplicar la normativa de declaración que le afecte. En el caso de las universidades, en concreto, pueden tener guías y normativas al respecto en, p.e., sus programas de doctorado o de máster. En cambio, si se trata de un artículo científico, lo razonable es atender a las guías que proporciona cada revista.

Un procedimiento general puede consistir en declarar el uso de la IA en el apartado de metodología o en una sección específica de "Uso de la IA" (a) detallando el uso concreto, (b) explicando como se aplicó la supervisión por parte del autor (p.e., con Veritas), y (c) facilitando datos de la IA, y si se juzga conveniente (d) complementando la declaración con datos de trazabilidad tales como reproducir los prompts y las respuestas literales de la IA en un anexo o como parte del *data set*. El lector interesado tiene algunas fuentes en las referencias de esta Tribuna.

7

8. Uso de la IA para esta Tribuna

Para no caer en el síndrome del herrero y su cuchillo de palo, va lo siguiente. Ninguna parte de los contenidos de este texto ha sido redactado por una IA. Todas y cada uno de los apartados de este artículo corresponde íntegramente a las muy limitadas capacidades intelectuales de este autor después de más de dos años de pelear con este tema. Ahora bien, utilizo de forma rutinaria sistemas de IA como Perplexity, Gemini, ChatGPT y Consensus como buscadores (además de Google Scholar, Scopus o WoS) como mis herramientas habituales de trabajo. Para contrastar las ideas principales de este trabajo, hice un prompt de comprobación de la coherencia de Veritas con el consenso científico usando Consensus y Undermind. Con ambos utilicé un prompt en el que preguntaba por el consenso a propósito de los puntos cardinales señalados más arriba. Me resultó de gran utilidad, pero en este caso para contrastar mis conjeturas, de paso me ayudó a descubrir algunos trabajos, p.e, los de Hosseini (ver referencias). El resultado de estos prompts pueden verlos a través de los enlaces que he añadido al final de las referencias en el apartado de materiales complementarios.

9. Conclusiones

Veritas es un marco de trabajo que permite utilizar sistemas de inteligencia artificial aplicando una transformación que es coherente tanto con los puntos cardinales de la IA como con las buenas prácticas y la ética habituales de los trabajos académicos.

Este marco está concebido pensando en trabajos académicos evaluativos, pero creo que en principio puede extrapolarse a cualquier producción característica de la comunicación académica si queremos aplicar estándares exigentes y a la vez queremos mitigar o anular la deuda cognitiva.

Corresponde señalar algunas cosas sobre el apartado de referencias. Se trata de las principales fuentes manejadas en los últimos meses por mi parte mientras contrastaba y refinaba Veritas. Veritas es muy modesto, pero me he asegurado que sea coherente lógicamente. Son por tanto, fuentes consultadas, no citadas. Dado el carácter de opinión y divulgación de esta Tribuna, creo que es el tratamiento adecuada en este caso.

Por último, me ha parecido útil añadir al acceso a las síntesis de la evidencia generadas por dos IA que considero de utilidad. He usado prompts que piden a las IA respectivas que evalúen los puntos cardinales. Las respuestas que han generado las IA respectivas

ofrecen a la vez interesantes síntesis narrativas y referencias muy relevantes analizadas. Sin duda son dos informes útiles para cualquier lector interesado.

10. Referencias consultadas y relacionadas

- Aguilera-Cora, Elisenda, Carlos Lopezosa, José Fernández-Cavia, & Lluís Codina. 2024. Accelerating research processes with Scopus AI: A place branding case study. *Revista Panamericana de Comunicación*, 6(1). <https://doi.org/10.21555/rpc.v6i1.3088>
- Bhavsar, Daivat, Laura Duffy, Hamin Jo, Cynthia Lokker, R. Brian Haynes, Alfonso Iorio, Ana Marušić, & Jeremy Y Ng (2024). Policies on artificial intelligence chatbots among academic publishers: a cross-sectional audit. *Research Integrity and Peer Review*, 10. <https://doi.org/10.1101/2024.06.19.24309148>
- Crompton, Helen, & Diane Burke. 2023. Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education*, 20(1). <https://doi.org/10.1186/s41239-023-00392-8>
- Font-Julián, Cristina I., Enrique Orduña-Malea, & Lluís Codina. 2024. ChatGPT Search as a tool for scholarly tasks: Evolution or return? *Infonomy*, 2(5). <https://doi.org/10.3145/infonomy.24.059>
- Gao, Runyang, Danghui Yu, Biao Gao, Heng Hua, Zhaoyang Hui, Jingquan Gao, & Cha Yin. 2025. Legal regulation of AI-assisted academic writing: challenges, frameworks, and pathways. *Frontiers in Artificial Intelligence*, 8. <https://doi.org/10.3389/frai.2025.1546064>
- Hosseini, Mohammad, David B. Resnik, & Kristi Holmes 2023. The ethics of disclosing the use of artificial intelligence tools in writing scholarly manuscripts. *Research ethics*, 19, 449-465. <https://doi.org/10.1177/17470161231180449>
- Hosseini, Mohammad, Lisa M. Rasmussen, & David B. Resnik. 2024. Using AI to write scholarly publications. *Accountability in Research*, 31(7), 715–723. <https://doi.org/10.1080/08989621.2023.2168535>
- Kasani, Payam H., Kee-Hyun Cho, Jae-Won Jang, & Cheol-Heui Yun. 2024. Influence of artificial intelligence and chatbots on research integrity and publication ethics. *Science Editing*. <https://doi.org/10.6087/kcse.323>
- Koplin, Julian. 2023. Plagiarism, academicethics, and the utilization of generative AI in academic writing. *International Journal of Applied Philosophy*. <https://doi.org/10.5840/ijap2023372202>
- Mondal, Himel, Shaikat Mondal, Joshil-Kumar Behera. 2024. Artificial intelligence in academic writing: Insights from journal publishers' guidelines. *Perspectives in Clinical Research*, 16, 56-57. https://doi.org/10.4103/picr.picr_67_24
- Penabad-Camacho, Liliana, María Morera-Castro & María-Amalia Penabad-Camacho. 2024. Guidelines for the use and reporting of Artificial Intelligence in scientific-academic journals. *Revista Electrónica Educare*, 28(2), 1-18. <https://doi.org/10.15359/ree.28-S.21028>
- Resnik, David B. , & Hosseini, Mohammad. 2024. The ethics of using artificial intelligence in scientific research: new guidance needed for a new tool. *Ai and Ethics*, 5, 1499–1521. <https://doi.org/10.1007/s43681-024-00493-8>
- Resnik, David B. , & Hosseini, Mohammad. 2025. Disclosing artificial intelligence use in scientific research and publication: When should disclosure be mandatory, optional, or unnecessary? *Accountability in Research*, 1–13. <https://doi.org/10.1080/08989621.2025.2481949>
- Torres-Salinas, Daniel, Wenceslao Arroyo-Machado. 2025. Enseñar e investigar con inteligencia artificial: una llamada a la reflexión. *BiD*, 54. <https://doi.org/10.1344/BID2025.54.01>
- UNESCO. 2023. *Guidance for generative AI in education and research*. <https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research>
- Veiga, Adèle (2025). Ethical guidelines for the use of generative artificial intelligence and artificial intelligence-assisted tools in scholarly publishing: a thematic analysis. *Science Editing*. <https://doi.org/10.6087/kcse.352>

Yousaf, Muhammad N. (2025). Practical considerations and ethical implications of using aArtificial intelligence in writing scientific manuscripts. *ACG Case Reports Journal*, 12. <https://doi.org/10.14309/crj.0000000000001629>

Materiales de referencia adicionales

- Carpeta drive con guías y normativa sobre el uso de la IA en trabajos académicos
- GAIDeT (Generative AI Delegation Taxonomy): A taxonomy for humans to delegate tasks to generative artificial intelligence in scientific research and publishing
- Síntesis de la evidencia generada con Consensus. https://consensus.app/search/evaluate-this-reasoning-if-there-is-a-consensus-on/50bZB47bQ5CMKcTXy4cc_Q/
- Síntesis de la evidencia generada con Undermind. <https://app.undermind.ai/report/6ba39c-de4e09765bc0ca85a0531ebd8ec5db440d55acdca4718c4f91f5d34573>