Revista científica editada por la Facultad de Información y Medios Audiovisuales de la Universidad de Barcelona
ISSN 1575-5886

BiD 55 (diciembre 2025)
DOI: 10.1344/bid2025.55.01

TRIBUNE BiD

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

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

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

Lluís Codina i
UniversitatPompeu Fabra
lluis.codina@upf.edu

Received: 03-12-2025
Accepted: 03-12-2025

Recommended citation

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

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.

Resumen

Presentación de los puntos cardinales del uso de la inteligencia artificial: las IA no pueden ser acreditas 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 l 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

Resum

Presentació dels punts cardinals de l'ús de la intel·ligència artificial: les IA no poden ser acredites com a autores; la responsabilitat d'un treball acadèmic és íntegra i exclusiva de l'autor; les persones autores han d'aplicar el pensament crític a l'ús de les IA i l'ús significatiu de la IA ha de ser declarat. Conseqüències daquests punts cardinals aplicats a lús 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. Introduction

As is well known, in just a few years, the use of generative artificial intelligence in academic work has become almost universal, or is on its way to becoming so. Before proceeding, for the purposes of what follows, I must clarify that by academic work I mean academic work that has an **evaluative character** .

To make it clear, I should point out that among the most representative examples of these productions are master's theses and doctoral dissertations. We can also easily add (quite the contrary) undergraduate theses, which often exhibit remarkable quality.

The important thing, especially in the case of doctoral theses, is that the highest quality academic works make very relevant contributions with both academic and social implications.

Furthermore, precisely because of the evaluative nature I mentioned earlier, successful completion of these programs leads to qualifications that enjoy widespread social recognition. Specifically, the final master's thesis results in a Master's degree, and doctoral dissertations in a Doctorate, thus opening up new professional opportunities for those who hold these qualifications.

An additional characteristic of this type of academic degree is that society expects them to be awarded fairly, that is, that the respective courts only grant them to those who have truly deserved them through their effort and not, for example, through the use of artificial intelligence replacing the author's effort.

What I will discuss here therefore refers to this type of academic production. However, I believe that much of what will be presented can be extrapolated to other formats of academic communication, such as scientific articles, since, albeit through different means, they also constitute evaluative work in some way. For example, in many countries, the publication of scientific articles leads to improvements in academic recruitment.

Finally, I would like to explain how this forum came about. For the past two years, I have been leading predoctoral training sessions and seminars for experts at my university (Pompeu Fabra), as well as at universities in Catalonia, Spain, and Latin America. The need for a practical way to teach and discuss the use of AI shaped the framework of procedures I call Veritas, which I present here. To arrive at this framework, I have spent these two years extensively researching the topic, using both conventional search engines and AI systems. Finally, the training experiences and discussions with colleagues in seminars have helped me refine and present Veritas in a way that I hope is both coherent and practical.

2. The cardinal points of AI ethics applied to academia

Those responsible for university degrees, from undergraduate to doctoral programs, are well aware that, to this day, there are numerous guides and guidelines on the use of artificial intelligence in academia, not only in evaluative work, but also (and above all) in research and scientific publication in general (see the References section).

And this is the point I wanted to make. Anyone who has examined some of these guidelines or regulations will realize that they are either abstract in nature or, when they go into detail, they leave areas that require interpretation.

Fortunately, the above does not preclude the emergence of a clear core of consensus from a careful reading of these documents. This emerging consensus is evident in at least the four cardinal points that will be presented (remember that a cardinal point is not necessarily a geographical location). Furthermore, these points generally overlap with the ethical guidelines of scientific journals.

I list them below in a way that I hope the reader interested in these topics will easily recognize:

- **No authorship** . AI cannot be credited as either author or co-author. This lack of authorship discourages citing AI as an authoritative source. Instead, citing the original source is encouraged.
- **Responsibility** . The responsibility for the content of an academic work rests entirely with the author.
- **Critical thinking** . Authors must use AI by applying their critical thinking skills to detect "hallucinations", avoid biases, and prevent errors or inconsistencies.
- **Transparency** . The significant use of AI in academic work must be declared in a transparent and traceable manner .

These four points are sometimes reduced to the first two, for good reason. The combination of the first and second, within the general framework of good scientific practice, naturally necessitates the third, and the fourth, for its part, is a requirement of general

academic tradition, since the use of AI does not invalidate conventional good practices. However, for the sake of clarity, I prefer to present all four points in their most explicit form.

3. The supply of artificial intelligence systems for academia

An apparent dilemma arises, for the reasons that will be seen, when we consider the characteristics of artificial intelligence systems designed for use in academic settings. I am referring to products, some well-known, such as the following four, which I will mention as an ostensive definition:

- Elicit
- SciSpace
- Consensus
- Undermind

This is artificial intelligence software that has clearly specialized in producing what we might call **evidence syntheses** (theoretical frameworks, literature reviews, etc.). From this starting point, they are geared (as their functions clearly demonstrate) towards generating content with a genuinely academic final format, such that a user *could* (note the conditional) directly use one or more sections of their work as part or in its entirety.

However, if we look back at the cardinal points mentioned, we must doubt that possibility. The reason is that when these cardinal points are applied to academic work, specifically points 1 and 2, but also point 4, a logical consequence emerges using both deduction and reductio ad absurdum. This consequence is as follows:

- No section of an evaluative academic paper can be written directly, either wholly or partially, by an artificial intelligence.

Let's see why: point 1, by prohibiting adding the name of an AI to the *byline,* logically, It also prohibits us from using AI content as if it were our own. The reason: if (a) that content is not from the authors, and (b) it cannot be attributed to the AI, then (c) it cannot be used as our own without committing plagiarism.

Point 2 acts as a reinforcement: given that (a) if we use an AI in some (admissible) phase of our work, and (b) in the event of any error or even any legal problem, only the (human) authors can be held responsible, it follows (c) that direct use, that is, without any kind of supervision or filtering, is logically discredited.

Point 4, for its part, completes the logical circle by requiring the use of critical thinking to verify, analyze, and, where necessary, filter the content generated by AI. And what about point 3? On the one hand, it's part of the demands that come with the new landscape of AI use, but at the same time, it's part of the traditional requirements of all academic work when it comes to providing the usual elements of transparency and traceability.

4. The problem

Given all of the above, and in the form of a question, the following problem arises: how can a pre- or postdoctoral researcher legitimately put the advantages of AI to use when applying it to academic work, especially if it is of an evaluative nature?

The way to address the problem is that, if AI-generated content is deemed useful, it must be **transformed** so that its application aligns with the expected ethics of evaluative work, which I have summarized in the four cardinal points. And this is where the proposed work or transformation, which I call Veritas, comes in.

5. Veritas

The function of Veritas is to provide a framework that allows the use of content generated by generative AI in academic work, while remaining compatible with the four cardinal points of AI ethics in academia. Veritas comprises four actions that, upon closer examination, are actually characteristics of standard academic best practices. They are as follows:

- Verify
- Assess
- Attribute
- Edit

The development of the elements of Veritas is as follows:

- **Verify the sources.** Any source that is a candidate for use and citation should be checked if it is considered relevant to the academic work. This verification includes all references of interest, whether to support ideas or any kind of data that we intend to use in our work.
- **Evaluate the constructs** (ideas, concepts, models, explanations, etc.) presented by the AI, using critical thinking. In this evaluation, the author (the doctoral candidate, for example) must not only ensure they grasp the essence of these constructs, but also relate them to the researcher's prior knowledge in their field, as well as to other elements of their theoretical or conceptual framework.
- **Attributing ideas from others** . Any valuable ideas or content that the researcher obtains from the two previous operations must be attributed to the original sources (the human authors, not the AI) to the extent that they consider incorporating them into their work. Attributions will be made either by explanation or paraphrase (plus the appropriate parenthetical citation) or by direct quotation (marking and delimiting the verbatim fragment plus the parenthetical citation).
- **Edit the content to ensure coherence** . Any content produced with AI assistance, except for direct quotes, must be edited by the author to generate a coherent discourse, unify the writing, and give the entire narrative of their work a distinct voice. This, in addition to providing the thesis with a coherent style, ensures that the author has taken responsibility and is able to explain all the concepts used in their work.

Now we can see in the following table the relationship between the cardinal points and the elements of Veritas (Table 1).

Table 1. Veritas and the cardinal points in the use of AI

Cardinal points	***Veritas Points***
AIs are not authors	Evaluation, Verification
AI cannot be cited as authors	Attribution
The author is fully responsible for the content	Evaluation, Verification, Editing
The author must use AI with critical thinking.	Evaluation, Verification

Source: own elaboration

6. Staggered critical interaction

The philosophy behind Veritas I call **staggered critical interaction (CSI)** . It's equivalent to the important principle called HITL, short for human-in-the- loop , but I think it's more expressive. HITL is a concept from AI engineering, and it maintains that a human is always actively involved in the operating cycle of an AI system to monitor, correct, and make key decisions.

CSI implies that the author doesn't simply use AI-generated content passively, or by "copying and pasting." In fact, it's the diametrical opposite of both. It's also the only sure way to mitigate or even eliminate **cognitive debt** . Recall that cognitive debt is generated when a human delegates advanced intellectual or cognitive functions to an AI. The main problem is that cognitive debt can lead to **cognitive atrophy** . This is hardly something a researcher in training (or even a senior researcher) would want.

The relationship between CSI and Veritas is as follows: the former (CSI) is the general philosophy that should govern the use of AI in academia (perhaps the use of AI in general); the latter (Veritas) is one of the operational forms in which it can be implemented. A comprehensive explanation of the meaning and implications of ICE is as follows:

- **Interaction**. This implies that the researcher does not passively accept the results of an AI. On the contrary, they interact with both the AI and its results. With the AI, they question its results and ask follow-up questions, redirect its responses, and so on. With the results, they use scrutiny.
- **Critique**. The above interaction must be guided by critical thinking, which obliges the researcher to apply, among other things, frameworks such as Veritas.
- **Staggered**. It is phased because it proceeds in stages, depending on the needs of using AI and the verification actions, etc., that must be carried out, the follow-up questions, etc. It also implies that the researcher is willing to invest intellectual effort in verification processes, etc.

7. Transparent and traceable declaration

We have the last of the cardinal points left, which states: "4. Transparency. The significant use of AI in an academic work must be declared in a transparent and traceable manner."

It's important to note here that this point doesn't stem from the use of AI, but rather from the standard best practices of any academic work, especially evaluative work where authors must disclose the methods they used to obtain their results. However, this disclosure alone adheres to the principle of transparency, which is necessary but not sufficient. When we use a methodology extensively, we must also provide traceability data.

At this point, it is the author's responsibility to apply the disclosure guidelines that apply to them. Universities, specifically, may have guidelines and regulations on this matter in, for example, their doctoral or master's programs. However, if it is a scientific article, it is reasonable to follow the guidelines provided by each journal.

A general procedure might consist of declaring the use of AI in the methodology section or in a dedicated "Use of AI" section by (a) detailing the specific use, (b) explaining how the author oversight was applied (e.g., with Veritas), and (c) providing AI data, and, if deemed appropriate, (d) supplementing the declaration with traceability data such as reproducing the AI prompts and verbatim responses in an appendix or as part of the *dataset* . Interested readers can find some sources in the references section of this article.

8. Use of AI for this Tribune

To avoid falling into the trap of the blacksmith with his wooden knife, let me clarify the following. None of the content in this text was written by an AI. Every single section of this article corresponds entirely to the very limited intellectual capabilities of this author after more than two years of grappling with this issue. However, I routinely use AI systems such as Perplexity, Gemini, ChatGPT, and Consensus as search engines (in addition to Google Scholar, Scopus, and Web of Science) as my standard working tools. To test the main ideas in this work, I created a prompt to check the consistency of Veritas with the scientific consensus using Consensus and Undermind. With both, I used a prompt asking about the consensus regarding the cardinal points mentioned above. I found it very useful, but in this case, to test my conjectures. It also helped me discover some other works, such as those by Hosseini (see references). The results of these prompts can be viewed through the links I have added at the end of the references in the supplementary materials section. Finally, the English version you are reading has benefited from a preliminary draft produced with AI assistance, which was later reviewed and refined.

9. Conclusions

Veritas is a framework that allows the use of artificial intelligence systems by applying a transformation that is consistent with both the cardinal points of AI and the usual good practices and ethics of academic work.

This framework is designed with evaluative academic work in mind, but I believe that in principle it can be extrapolated to any production characteristic of academic communication if we want to apply demanding standards and at the same time mitigate or eliminate cognitive debt.

It's worth noting a few things about the references section. These are the main sources I've used in recent months while verifying and refining Veritas. Veritas is quite modest, but I've ensured its logical consistency. Therefore, these are sources consulted, not cited. Given the opinion-based and informative nature of this column, I believe this is the appropriate approach in this case.

Finally, I found it useful to add access to the evidence summaries generated by two AIs that I consider helpful. I used prompts that ask the respective AIs to evaluate cardinal points. The responses generated by the AIs offer both interesting narrative summaries and highly relevant analyzed references. These are undoubtedly two useful reports for any interested reader.

10. References consulted and related

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, Adéle (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

Additional reference materials

- Drive folder with guides and regulations on the use of AI in academic work
- GAIDeT (Generative AI Delegation Taxonomy): A taxonomy for humans to delegate tasks to generate artificial intelligence in scientific research and publishing
- Summary of the evidence generated with Consensus. https://consensus.app/search/evaluate-this-reasoning-if-there-is-a-consensus-on/50bZB47bQ5CMKcTXy4cc_Q/
- Summary of the evidence generated with Undermind. https://app.undermind.ai/report/6ba39cde4e09765bc0ca85a0531ebd8ec5db440d55acdca4718c4f91f5d34573