

Disclosing generative AI use for writing assistance should be voluntary

Research Ethics

1–8

© The Author(s) 2025

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/17470161251345499

journals.sagepub.com/home/rea

Mohammad Hosseini^{1,2}, Bert Gordijn³,
Gregory E. Kaebnick⁴ and Kristi Holmes^{1,2}

Abstract

Researchers have been using generative artificial intelligence (GenAI) to support writing manuscripts for several years now. However, as GenAI evolves and scientists are using it more frequently, the case for mandatory disclosure of GenAI for writing assistance continues to diverge from the initial justifications for disclosure, namely (1) preventing researchers from taking credit for work done by machines; (2) enabling other researchers to critically evaluate a manuscript and its specific claims; and (3) helping editors determine if a submission satisfies their editorial policies. Our initial position (communicated through previous publications) regarding GenAI use for writing assistance was in favor of mandatory disclosure. Nevertheless, as we show in this paper, we have changed our position and now support instituting a voluntary disclosure policy because currently (1) the credit due to machines for assisting researchers is moving below the threshold of requiring recognition; (2) it is impractical (if not impossible) to accurately specify what parts of the text are human-/GenAI-generated; and (3) disclosures could increase biases against non-native speakers of the English language and compromise the integrity of the peer review system. Consequently, we argue, it should be up to the authors of manuscripts to disclose their use of GenAI for writing assistance. For example, in disciplines where writing is the hallmark of originality, or when authors believe disclosure is beneficial, a voluntary checkbox in manuscript submission systems, visible only after publication (rather than a free-text note in the manuscripts) would be preferable.

¹Northwestern University Feinberg School of Medicine, Chicago, IL, USA

²Galter Health Sciences Library and Learning Center, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

³Dublin City University, Ireland

⁴The Hastings Center, Philipstown, NY, USA

Corresponding author:

Mohammad Hosseini, Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, 680 N Lake Shore Dr, Suite 1400, Chicago, IL 60611-3008, USA.

Email: mohammad.hosseini@northwestern.edu



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

Keywords

artificial intelligence, disclosure, editorial policies, publication ethics, peer review, writing

Introduction

Reporting of tools, methods and individual contributions in a research project supports research integrity by upholding principles of honesty, transparency, and reproducibility (Leonelli, 2023). To the extent that disclosures are aligned with these principles, they also impact the public's views about science and particularly its trustworthiness (Shamoo and Resnik, 2022). With the advent of generative artificial intelligence (GenAI) and its increased use in various scholarly tasks, researchers are expected to apply the same ethical principles and disclose their use of these tools. Arguments in favor of disclosure of GenAI hold that it:

- (1) Prevents researchers from taking credit for work done by machines (Verhoeven et al., 2023);
- (2) Enables other researchers to critically evaluate a manuscript and its specific claims (Hosseini et al., 2023a); and
- (3) Helps editors determine if a submission satisfies their editorial policies since different journals have widely dissimilar policies on the use and disclosure of GenAI (Kaeznick et al., 2023).

Here we show that these three arguments do not provide sufficient justification for mandating disclosure when GenAI is used for writing assistance (i.e. using GenAI to revise existing text to improve grammar, style, and readability). Accordingly, we suggest that disclosing GenAI use for writing assistance should be voluntary. Distinguishing between writing assistance and writing entire sentences/paragraphs from scratch is important because GenAI is trained on existing data. When GenAI generates text from scratch, it only draws from previously published text, thereby generating content that may not be original. This lack of originality and text regurgitation negatively impacts on the integrity of research because it makes attributions of credit to past authors extremely more complex and obfuscates the sources of ideas. While one could argue that the difference between writing assistance and text creation is gradual, and that the former could gradually morph into the latter, distinguishing between the two is still possible. This is because GenAI delivers what users request. If users ask GenAI to improve grammar, style, and readability, that will be the generated outcome, but if they ask GenAI to create new sentences or expand existing text, that is what they will get.¹

This position (that disclosing the use of GenAI for writing assistance should be voluntary) departs from views previously endorsed by authors of this work, where we held that disclosure should be mandatory (Hosseini et al., 2023a; Kaeznick et al., 2023). The next three sections address the three justifications mentioned above for disclosure of GenAI. In each section, we demonstrate that, although there are general arguments in favor of GenAI disclosure, these arguments do not hold in the specific case of using GenAI for writing assistance. For this case, we argue, *mandatory* disclosure policies are unnecessary, lead to tensions, and are counterproductive. We conclude by arguing that disclosure should be voluntary when GenAI is used for writing assistance and offer recommendations to implement this policy.

Disclosure to prevent researchers from taking credit for work done by machines

The ethical principle “give credit where credit is due” requires giving authorship credit to those who satisfy the required criteria for authorship, and clarifying who has done what, for example, in the contributorship statements or using the CRediT taxonomy (Vasilevsky et al., 2021). The recent increase in the use of GenAI for writing has intensified the debate about authorship, especially about how to acknowledge the role of non-human contributors in writing. In the early days of ChatGPT and in the absence of specific policies on using GenAI in writing, some researchers attributed authorship credit to these systems. This trend was quickly stifled by guidelines that banned naming GenAI as authors and suggested other means to disclose the use of GenAI in writing (Committee on Publication Ethics (COPE), 2023). It should be noted that some scholars still argue in favor of naming GenAI as authors (Abernethy, 2024), but we consider such claims questionable because of the significance of responsibilities and accountabilities in considerations about authorship (Kaeznick et al., 2023).

Although there is currently a near-widespread consensus in the scholarly community that GenAI systems cannot be considered authors, their ubiquity and growing integration into writing tools (e.g. MS Word) will likely make human and GenAI contributions increasingly intertwined. As a result, the justification for requiring disclosure (to prevent researchers from taking credit for writing assistance done by machines) might eventually become less relevant because

- (a) A clear demarcation between human and GenAI contributions will be impractical (if not impossible). Our current use of unique tools incorporated into word processors (e.g. dictionary, spelling checker, thesaurus, reference manager) demonstrates that when writing, it is impractical to keep track of what tools we used and when. This is because writing is a task that requires focus and attention, and constant documentation and disclosure of all used tools and their impact on each sentence is painstakingly cumbersome. Researchers have not done this for dictionaries or thesauruses and are unlikely to do it for GenAI.
- (b) As writing tools become more ubiquitous, the credit due to them for assisting researchers could move below the threshold of requiring receiving recognition. As GenAI tools are increasingly integrated into word processors, they will likely be moved below this threshold, and regarded as tools that, despite their indispensability in the writing process, do not require acknowledgment. Even before the advent of GenAI, researchers relied on other systems that were essential for writing but were never formally acknowledged in publications because of their ubiquity. For instance, with the increased accessibility of computers and the dominance of digital systems in the publication ecosystem, word processors became indispensable for drafting and submitting papers. Yet, we do not acknowledge these systems because they are so ubiquitous that they are considered part of the infrastructure (like the role of a post office in a research project that involves collecting samples – indispensable but unnecessary to disclose or acknowledge).

Nevertheless, one may argue that GenAI offers way more assistance than previous tools like dictionaries, and therefore, disclosure should still be mandatory. While we agree that GenAI is much more powerful than a dictionary, our response to this argument is a pragmatic one: If the purpose of mandating disclosure is to increase transparency, this can only succeed if we can enforce and police the mandate. However, currently we cannot enforce and police such mandates simply because detecting GenAI-generated text is not always possible, and consequently, any

policy to that effect is superfluous. Disclosure policies released by some publishers suggest that the imminent collapse of such mandates is recognized. For example, Springer does not require disclosure of AI-assisted copyediting (Springer, 2023). The editorial policies of their flagship journal, *Nature*, state:

The use of an LLM (or other AI-tool) for “AI assisted copy editing” purposes does not need to be declared. In this context, we define the term “AI assisted copy editing” as AI-assisted improvements to human-generated texts for readability and style, and to ensure that the texts are free of errors in grammar, spelling, punctuation and tone. These AI-assisted improvements may include wording and formatting changes to the texts, but do not include generative editorial work and autonomous content creation (Nature Portfolio, 2023).

Disclosure to enable evaluation of a manuscript and its specific claims

For those reading a scholarly manuscript, accurate disclosure of GenAI use for writing tasks (e.g. what parts were written or edited by GenAI) will link specific parts of a manuscript to machines and will enable better evaluation of its specific claims (Hosseini et al., 2023b). The rationale for this demand for disclosure is that, if readers know what GenAI system and prompts were used to write a paragraph, they can read sections affected by GenAI with caution and be mindful about its possible limitations (Tang et al., 2024). Nevertheless, for disclosures to achieve such goals, they must be extremely specific and detailed (e.g. to include the used prompt, used GenAI, the sentence/paragraph that was written/revised by GenAI etc.), which is cumbersome (if not impractical). As human and Gen-AI contributions get increasingly intertwined, the expectation to disclose GenAI use after each impacted sentence/paragraph seems excessive and could be a deterrent for transparency. Mandatory disclosures are particularly burdensome because GenAI disclosures are not currently automated and require manual effort. Consequently, it is highly likely that disclosure may take the form of a *blanket statement* in the methods or acknowledgments section, along the lines of “ChatGPT was used to assist writing this manuscript” or “ChatGPT was used to improve the readability of the text.”

Without any further clarification about what parts of the text were affected (to enable reviewers to distinguish human- and machine-written sections), and who used GenAI (to enable attribution of accountabilities), the impacts of GenAI disclosure are not aligned with the envisioned purpose (of enabling the reader to get an *accurate* understanding of the writing of the paper). Exploring disclosure policies published by authoritative organizations suggests that these misalignments are well recognized. For example, the Committee of Publication Ethics (COPE) position statement on the use of AI in research publications does not *explicitly* demand specification of sections that were co-written by AI:

Authors who use AI tools in the writing of a manuscript, production of images or graphical elements of the paper, or in the collection and analysis of data, must be transparent in disclosing in the Materials and Methods (or similar section) of the paper how the AI tool was used and which tool was used (COPE, 2023).

The same goes for the International Committee of Medical Journal Editors’ (ICMJE) Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly work in Medical Journals:

Authors who use such technology should describe, in both the cover letter and the submitted work in the appropriate section if applicable, how they used it. . . . For example, if AI was used for writing assistance, describe this in the acknowledgment section. If AI was used for data collection, analysis, or figure generation, authors should describe this use in the methods (ICMJE, 2024).

In their guidelines about Chatbots, Generative AI, and Scholarly Manuscripts, the World Association of Medical Journal Editors (WAME) are slightly more specific, but again, they are satisfied with blanket statements with the caveat that they also demand disclosure of used prompts:

Authors should be transparent when chatbots are used and provide information about how they were used. The extent and type of use of chatbots in journal publications should be indicated. . . . Authors submitting a paper in which a chatbot/AI was used to draft new text should note such use in the acknowledgment; all prompts used to generate new text, or to convert text or text prompts into tables or illustrations, should be specified (Zielinski et al., 2023).

Comparing these guidelines demonstrates another issue, namely that there are different views about *how* and *where* in a manuscript (or elsewhere like a cover letter) disclosure should happen and *what* should be disclosed. In the absence of harmonized disclosure policies, we are likely to see inconsistent disclosure practices. If the research ethics and integrity experts have learned one thing about lack of harmony in policies is that it ultimately deters compliance and causes more problems overall. The application of different and inconsistent criteria for authorship across guidelines (with implications for who should be an author), varying definitions of misconduct in different countries (with implications for what counts as misconduct for misdemeanors other than falsification, fabrication, and plagiarism) and disparities in terms of function and power of research integrity officers in different countries (with implications for dispute resolutions) are illustrative examples of what lack of harmony can lead to (Desmond and Dierickx, 2021; Hosseini and Lewis, 2020; Videnova et al., 2024).

Either way, without any clarification about *what parts* of a text are written by GenAI assistance and *who* has used it, disclosures diverge from the envisioned purpose of enabling an accurate evaluation of a manuscript. Since such detailed disclosures are extremely burdensome (especially in light of increased human-machine interaction) and cannot be policed, moving toward a voluntary disclosure approach is more practical.

Disclosure to help editors determine if a submission satisfies editorial policies

Different journals have adopted dissimilar policies regarding GenAI use.² When considering the impacts of transparent disclosure of GenAI use for writing assistance on the editorial process, we should explore how editorial assessment and decision making will be influenced by disclosures. If certain GenAI uses are banned in journals (e.g. *Science* does not accept GenAI images), *truthful* disclosure statements enable editors to identify whether submitted manuscripts comply with policies, and thus, are necessary. However, using GenAI for writing assistance has not been banned by any journal. Accordingly, disclosures would not serve a purpose in terms of ensuring adherence to policy beyond the fact that some journals (such as *Science*) have mandated disclosure (which as mentioned in the first section, cannot be policed):

“Authors who use AI-assisted technologies as components of their research study or as aids in the writing or presentation of the manuscript should note this in the cover letter and in the acknowledgments section of the manuscript” (Thorp and Vinson, 2023).

Additionally, disclosing the use of GenAI in ways that would be visible to editors and peer reviewers could compromise the integrity of the double-blind peer review process. Whether disclosure happens in the body of the text, acknowledgments section, or in the cover letter, editors’ and peer-reviewers’ awareness of GenAI use for writing assistance can have a negative impact on their editorial decisions and should only be visible after the publication of a manuscript. Especially given reported biases toward non-native English authors (Amano et al., 2023; Hadan et al., 2024; Smith et al., 2023), if disclosing the use of GenAI for writing assistance prompts editors/reviewers that the author(s) might have English as a second language, disclosures may inadvertently affect how a manuscript is perceived and reviewed. This is particularly problematic because one of the main arguments in favor of using GenAI in writing tasks is that it levels the playing field for those working and writing in a second language (Resnik and Hosseini, 2023).

Toward a voluntary disclosure approach

In this paper we showed that, although there are general arguments in favor of disclosing GenAI use in research, these arguments do not hold in the specific case of using GenAI for writing assistance. Accordingly, we suggest that disclosing GenAI use for writing assistance should be voluntary. This position aligns with the claim that using AI “to edit existing text for grammar, spelling or organization” does not constitute substantial use, and that disclosure should be optional (Resnik and Hosseini, 2025: 7).

We acknowledge that some may perceive the involvement of GenAI in writing as a potential limitation, for example, raising concerns about originality, intellectual contribution, or the accuracy of the writing (Rentier, 2024). This limitation may be especially relevant to disciplines where human authorship and novelty of expressions are highly valued. To address this issue, we first made a distinction between using GenAI to generate new text versus using it to revise existing text (i.e. enhancing grammar, readability, and style). Second, we argued that it should be up to the authors of manuscripts to disclose their use of GenAI for writing assistance, and generated text should always be checked for accuracy, precision and relevance. For example, in disciplines where writing is the hallmark of originality, or when authors or others involved need to avail of a platform to demonstrate human versus machine contributions, this can still be facilitated without instituting a blanket mandatory disclosure policy (which is unenforceable as well).

Toward this end, and to better support voluntary disclosures for contexts where disclosure may be desirable by authors, we propose implementing a disclosure mechanism at the point of manuscript submission by means of machine-readable checkboxes (instead of free-text disclosure in the body of the manuscript). Researchers could voluntarily indicate whether GenAI was used in the writing process, and for this purpose, a taxonomy similar to the CRediT taxonomy of contributions can be developed to highlight specific uses of GenAI in writing assistance. Examples could include Grammar Check, Style Enhancement, Content Summarization, and Finding Examples to Improve Comprehension. However, to prevent any bias during the peer review and editorial process, this information should only become available once the article is accepted and published. Upon publication, such details could be disclosed as part of declarations and registered as metadata, to provide the desired transparency without influencing the initial assessment of the work. This approach would not only help maintain the integrity of the peer-review process but also ensure that, in cases where disclosures are desirable by authors, the presence and role of GenAI in academic writing are clearly documented.

Acknowledgements

We are grateful for the feedback provided by two anonymous peer reviewers. MH and GEK presented an early draft of this work during the 2024 annual conference of the American Society for Bioethics and Humanities in St. Louis.

ORCID iDs

Mohammad Hosseini  <https://orcid.org/0000-0002-2385-985X>

Bert Gordijn  <https://orcid.org/0000-0002-3686-8659>

Kristi Holmes  <https://orcid.org/0000-0001-8420-5254>

Ethical considerations

Ethical approval is not relevant to this study because no human subjects were used.

Author contributions

All authors have contributed extensively to the work presented in this paper and satisfy the criteria for authorship. Individual contributions, described by CRediT roles include. MH: Conceptualization, Investigation, Methodology, Project administration, Writing—original draft, Writing—review & editing. BG: Investigation, Validation, Writing—review & editing. GEK: Investigation, Validation, Writing—review & editing. KH: Funding acquisition, Investigation, Supervision, Validation, Writing—original draft, Writing—review & editing.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Institutes of Health's National Center for Advancing Translational Sciences (UM1TR005121).

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Notes

1. It should be noted that biases and errors could still creep in, and it is the responsibility of authors to check the revised text for accuracy, precision and relevance. As shown in an example offered by one of the anonymous peer reviewers: “For instance, imagine requesting GenAI to improve and shorten the following sentence: ‘we didn’t find any difference between male and female participants,’ and getting as a result: ‘no difference was found across genders.’ Now, here is a shift from sex to gender that the human user might or might not endorse.”
2. For example, Science outright bans any use of GenAI for image generation and specifically notes that “Editors may decline to move forward with manuscripts if AI is used inappropriately” (Thorp and Vinson, 2023). Nature, on the other hand, highlights several exceptions for the use of GenAI images (Nature Portfolio, 2023).

References

- Abernethy NJ (2024) Let stochastic parrots squawk: Why academic journals should allow large language models to coauthor articles. *AI and Ethics*. Epub ahead of print 19 September 2024. DOI: 10.1007/s43681-024-00575-7.

- Amano T, Ramírez-Castañeda V, Berdejo-Espinola V, et al. (2023) The manifold costs of being a non-native English speaker in science. *PLoS Biology* 21(7): e3002184.
- Committee on Publication Ethics (COPE) (2023) Authorship and AI tools. Available at: <https://publication-ethics.org/guidance/cope-position/authorship-and-ai-tools> (accessed 21 January 2025).
- Desmond H and Dierickx K (2021) Research integrity codes of conduct in Europe: Understanding the divergences. *Bioethics* 35(5): 414–428.
- Hadan H, Wang DM, Mogavi RH, et al. (2024) The great AI witch hunt: Reviewers' perception and (mis) conception of generative AI in research writing. *Computers in Human Behavior: Artificial Humans* 2(2): 100095.
- Hosseini M, Resnik DB and Holmes K (2023a) The ethics of disclosing the use of artificial intelligence tools in writing scholarly manuscripts. *Research Ethics* 19(4): 449–465.
- Hosseini M, Rasmussen LM and Resnik DB (2023b) Using AI to write scholarly publications. *Accountability in Research* 31(7): 715–719.
- Hosseini M and Lewis J (2020) The norms of authorship credit: Challenging the definition of authorship in the European Code of Conduct for Research Integrity. *Accountability in Research* 27(2): 80–98.
- International Committee of Medical Journal Editors (ICMJE) (2024) Recommendations | Defining the role of authors and contributors. Available at: <https://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html> (accessed 4 October 2024).
- Kaebnick GE, Magnus DC, Kao A, et al. (2023) Editors' statement on the responsible use of Generative AI Technologies in Scholarly Journal Publishing. *Hastings Center Report* 53(5): 3–6.
- Leonelli S (2023) *Philosophy of Open Science*. Cambridge: Cambridge University Press. Available at: <https://www.cambridge.org/core/elements/philosophy-of-open-science/0D049ECF635F3B676C03C6868873E406> (accessed 21 January 2025).
- Nature Portfolio (2023) Artificial intelligence (AI) policy. Available at: <https://www.nature.com/nature-portfolio/editorial-policies/ai> (accessed 21 January 2025).
- Rentier ES (2024) To use or not to use: exploring the ethical implications of using generative AI in academic writing. *AI and Ethics*. Epub ahead of print 20 December 2024. DOI: 10.1007/s43681-024-00649-6.
- Resnik DB and Hosseini M (2023) The impact of AUTOGEN and similar fine-tuned large language models on the integrity of scholarly writing. *American Journal of Bioethics* 23(10): 50–52.
- Resnik DB and Hosseini M (2025) Disclosing artificial intelligence use in scientific research and publication: When should disclosure be mandatory, optional, or unnecessary? *Accountability in Research* 1–13.
- Shamoo AE and Resnik DB (2022) *Responsible Conduct of Research*, 4th edn. New York, NY: Oxford University Press.
- Smith OM, Davis KL, Pizza RB, et al. (2023) Peer review perpetuates barriers for historically excluded groups. *Nature Ecology & Evolution* 7(4): 512–523.
- Springer (2023) Artificial intelligence policy. Available at: <https://www.springer.com/gp/editorial-policies/artificial-intelligence-ai-/25428500> (accessed 14 January 2025).
- Tang A, Li K-K, Kwok KO, et al. (2024) The importance of transparency: Declaring the use of generative artificial intelligence (AI) in academic writing. *Journal of Nursing Scholarship* 56(2): 314–318.
- Thorp HH and Vinson V (2023) Change to policy on the use of generative AI and large language models. *Science Editor's Blog*. 16 November 2023. Available at: <https://www.science.org/content/blog-post/change-policy-use-generative-ai-and-large-language-models> (accessed 21 January 2025).
- Vasilevsky NA, Hosseini M, Teplitzky S, et al. (2021) Is authorship sufficient for today's collaborative research? A call for contributor roles. *Accountability in Research* 28(1): 23–43.
- Verhoeven F, Wendling D and Prati C (2023) ChatGPT: When artificial intelligence replaces the rheumatologist in medical writing. *Annals of the Rheumatic Diseases* 82(8): 1015–1017.
- Videnoja K, Tauginienė L and Löfström E (2024) Family without kinship – The pluralism of European regulatory research integrity systems and its implications. *Accountability in Research*. Epub ahead of print 24 April 2024. DOI: 10.1080/08989621.2024.2345710.
- Zielinski C, Winker M, Aggarwal R, et al. (2023) WAME recommendations on chatbots and generative artificial intelligence in relation to scholarly publications. Available at: <https://wame.org/page3.php?id=106> (accessed 21 January 2025).