Authorship, Plagiarism, and Generative AI in Mathematical Research

Jordi-Lluís Figueras
Department of Mathematics, Uppsala University
Box 480, 751 06 Uppsala, Sweden
figueras@math.uu.se
ORCID: 0000-0002-0535-4137

March 16, 2026

Abstract

This article aims to offer some thoughts on whether and how generative AI (genAI) should be used in mathematical research (but also applies to learning and teaching). The main purpose is to organize some of the very basic ideas and questions around it. In particular, we construct the question of good and bad practice of genAI use based on the question of authorship. For doing so, we start from the concept of authorship and its relation to plagiarism. We list the main reasons why plagiarism is considered wrong, and good and bad practices for using AI in the generation and writing of academic texts. We finish with a small reflection on the value of heavily AI-generated texts. We do not intend this text to be exhaustive; take it as a starting point.

1 Introduction

Recently, the debate on the use of AI tools for generating mathematical results has intensified considerably. It seems that genAI is becoming more capable of helping to tackle research problems: the advances are broad and lie in different aspects such as code generation, proof formalization, or even progress on some open problems; see for example [19, 9, 22, 7, 17]. This has raised several reactions within the community. One of them concerns what the goal of mathematics is. In [1], the authors raise the following key point: the objective is not only to produce formally correct statements, but also to develop understanding, identify good questions, and build useful conceptual frameworks. An interesting initiative reflecting this trend is the upcoming workshop *Machine Learning for Mathematics* at the Simons Laufer Mathematical Sciences Institute (SLMath) in Spring 2027.

However, there are legitimate questions (see [1, 14]) about the use of genAI in mathematical research. Because of all this, in this note we try to shed light on the question of *when the use of (generative) AI poses a threat to authorship.*

We think that the central question that one needs to answer when using genAI tools (or any other) is:

*Am I **retaining the authorship** of the work if I do X?*

Just by using this question alone we get a good enough moral compass that help us evaluating what uses of genAI make us loosing the authorship. [1] We get deep into all this by first briefly exploring the concepts of authorship and plagiarism; and then seeing the moral implications of the latter. Then, we can apply this moral compass to analyse which of the uses of genAI makes us retain authorship.

2 Authorship, Plagiarism, and the Responsible Use of Generative AI

The ICMJE (International Committee of Medical Journal Editors) provides an appealing definition of authorship in [10]. It recommends that authorship should be based on the following four criteria:

1. Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; AND
2. Drafting the work or reviewing it critically for important intellectual content; AND
3. Final approval of the version to be published; AND
4. Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

This provides a practical working definition of authorship. Under these criteria, failing to satisfy any of the points weakens authorship claims. One common way this can happen is through plagiarism.

In [21] plagiarism is defined as

> Plagiarism is ‘‘the use of others’ published and unpublished ideas or words (or other intellectual property) without attribution or permission.’’

We care about plagiarism because of ethics and morality: [8] defines it as

> Morality can be used ‘‘descriptively to refer to certain codes of conduct endorsed by a society or a group.’’

[1] Sometimes not retaining authorship is a very valid option (as long as we do not intend to falsely claim it).

If academic communities depend on shared norms of honesty, attribution, and accountability, then plagiarism is not just a technical citation error but a moral violation of those norms.

Core values related to plagiarism and attribution are emphasized across research-integrity and publication-ethics frameworks and editorials [2, 13, 4, 5, 14, 18, 6]. They include:

1. **Integrity and Authenticity:** Plagiarism misrepresents who produced the work and undermines the honesty required for genuine scholarship.
2. **Respect for Intellectual Property:** It appropriates others' ideas or expressions without proper credit, violating ownership and creative rights.
3. **Transparency and Attribution:** Missing citations hide the origin of claims and prevent readers from verifying or building on prior work.
4. **Fairness and Equity:** Copying gives an unearned advantage over people who did original work and followed the rules.
5. **Trust and Credibility:** Once plagiarism is discovered, confidence in the author—and often in the institution—drops sharply.
6. **Innovation and Progress:** Rewarding imitation over originality discourages new ideas and slows collective intellectual advancement.
7. **Hindrance to Learning and Personal Growth:** When people outsource core thinking and writing tasks, they lose opportunities to develop understanding, judgment, and independent academic voice.

Based on these points, unattributed, substantial use of genAI can constitute plagiarism and weaken authorship claims; here, "substantial use" means delegating originality-bearing text or code to the model without transparent attribution and without critical human intellectual contribution.

We should not conclude that all use of genAI for document or code creation must be avoided. Generative AI can be a useful tool in many processes. Some of them (see for example [20, 15, 16]) are: [2]

- explore the proof of results and/or sharpen hypotheses,
- check grammar and spelling,
- ask for feedback about the text or the arguments in the text (you can even prompt it to act as a referee or editor from a journal),
- search for references and sources,
- summarize notes and papers,
- propose counterarguments to stress-test your claims,

[2] We do not keep in this list one very important use of genAI since it does not have relation to authorship. It is *holding a Socratic dialogue to understand concepts*.

- translate your text into other languages,
- ask for help with reformatting, and
- propose small refactorings while keeping full review responsibility.

Practical prompting techniques (such as clear task instructions, few-shot exemplars, and prompt pattern selection) can improve outcomes, but they also carry risks like sensitivity to phrasing and hidden bias; see the systematic survey in [12]. Moreover, one must not forget to acknowledge the use of genAI explicitly and explain how and where it was used. This is aligned with the EMS Code of Practice for Mathematical Publication [5], which asks authors to disclose AI or other automated tools and the nature of their use, while keeping full human responsibility for the final content; it also requires referees to disclose AI-assisted drafting of reports to editors and take full responsibility for those reports.

In addition to the previous list, we can include some "not-to-do" items:

- don't upload confidential, unpublished, or sensitive material (especially confidential drafts, notes, or preprints from colleagues),
- don't trust AI output blindly; always verify important facts,
- don't copy-paste AI text without editing for accuracy and tone,
- don't plagiarize or submit AI work as entirely your own,
- don't let AI replace your own judgment and critical thinking.

Moreover, under some conditions, genAI may reproduce training text verbatim, so blindly copy-pasting generated output can create plagiarism risk [3, 11].

Finally, *what about large chunks of text solely produced by genAI if they are great and meaningful?* If the text is relevant and useful, that is great! The mathematics community needs texts that help us advance in understanding. But writing a short and generic prompt for genAI does not make you the author of the output. That is because you did not **substantially contribute to the conception or design of the work**.

3 Funding and acknowledgments

The author thanks Joaquim Puig and Alex Haro for their fruitful comments. The author has been partially supported by VR Grant 2024-04764.

ChatGPT 5.2 and OpenAI Codex (gpt-5.3-codex) were used for formatting, reference search, grammar and spelling checks, and structural feedback; the author remains the sole author and fully responsible for the final content.

References

[1] Mohammed Abouzaid, Andrew J. Blumberg, Martin Hairer, Joe Kileel, Tamara Kolda, Paul Nelson, Daniel Spielman, Nikhil Srivastava, Rachel Ward, Shmuel Weinberger, and Lauren Williams. First Proof. arXiv preprint: https://arxiv.org/abs/2602.05192, February 2026. Submitted February 5, 2026; accessed March 2, 2026. URL: https://arxiv.org/abs/2602.05192.

[2] ALLEA - All European Academies. The European Code of Conduct for Research Integrity: Revised Edition 2023. DOI landing page: https://doi.org/10.26356/ECOC, June 2023. Accessed March 2, 2026. doi:10.26356/ECOC.

[3] Nicholas Carlini, Florian Tramer, Eric Wallace, Matthew Jagielski, Ariel Herbert-Voss, Katherine Lee, Adam Roberts, Tom Brown, Dawn Song, Ulfar Erlingsson, Alina Oprea, and Colin Raffel. Extracting Training Data from Large Language Models. In *30th USENIX Security Symposium (USENIX Security 21)*, pages 2633–2650, 2021. Accessed March 2, 2026. URL: https://www.usenix.org/conference/usenixsecurity21/presentation/carlini-extracting.

[4] Committee on Publication Ethics (COPE) and Directory of Open Access Journals (DOAJ) and Open Access Scholarly Publishers Association (OASPA) and World Association of Medical Editors (WAME). Principles of Transparency and Best Practice in Scholarly Publishing. Version 4, website: https://doi.org/10.24318/cope.2019.1.12, September 2022. Version 4 published September 15, 2022; accessed March 2, 2026. doi:10.24318/cope.2019.1.12.

[5] European Mathematical Society. Code of Practice for Mathematical Publication. Website: https://euromathsoc.org/code-of-practice, March 2025. Approved March 19, 2025; accessed March 2, 2026. URL: https://euromathsoc.org/code-of-practice.

[6] Annette Flanagin, Kirsten Bibbins-Domingo, Michael Berkwits, and Stacy L. Christiansen. Reporting the Use of AI in Research and Scholarly Publication. *JAMA*, 331(13):1096–1098, April 2024. Accessed March 2, 2026. doi:10.1001/jama.2024.3692.

[7] Bogdan Georgiev, Javier Gómez-Serrano, Terence Tao, and Adam Zsolt Wagner. Mathematical exploration and discovery at scale. arXiv preprint: https://arxiv.org/abs/2511.02864, November 2025. First submitted November 3, 2025; last revised December 22, 2025 (v3); accessed March 2, 2026. URL: https://arxiv.org/abs/2511.02864, doi:10.48550/arXiv.2511.02864.

[8] Joshua Gert and Bernard Gert. The Definition of Morality. In Edward N. Zalta and Uri Nodelman, editors, *The Stanford Encyclopedia*

of Philosophy. Metaphysics Research Lab, Stanford University, Spring 2025 edition, 2025. Archived edition recommended for citation; accessed March 2, 2026. URL: https://plato.stanford.edu/archives/spr2025/entries/morality-definition/.

[9] Thomas Hubert, Rishi Mehta, Laurent Sartran, Miklós Z. Horváth, Goran Žužić, Eric Wieser, Aja Huang, Julian Schrittwieser, Yannick Schroecker, Hussain Masoom, Ottavia Bertolli, Tom Zahavy, Amol Mandhane, Jessica Yung, Iuliya Beloshapka, Borja Ibarz, Vivek Veeriah, Lei Yu, Oliver Nash, Paul Lezeau, Salvatore Mercuri, Calle Sönne, Bhavik Mehta, Alex Davies, Daniel Zheng, Fabian Pedregosa, Yin Li, Ingrid von Glehn, Mark Rowland, Samuel Albanie, Ameya Velingker, Simon Schmitt, Edward Lockhart, Edward Hughes, Henryk Michalewski, Nicolas Sonnerat, Demis Hassabis, Pushmeet Kohli, and David Silver. Olympiad-level formal mathematical reasoning with reinforcement learning. *Nature*, November 2025. Published online November 12, 2025; accessed March 2, 2026. doi:10.1038/s41586-025-09833-y.

[10] International Committee of Medical Journal Editors. Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals. Website: https://www.icmje.org/icmje-recommendations.pdf, January 2026. Updated January 2026, accessed March 2, 2026. URL: https://www.icmje.org/icmje-recommendations.pdf.

[11] Kaixin Lan, Tao Fang, Derek Wong, Yabo Xu, Lidia Chao, and Cecilia Zhao. FOCUS: Forging originality through contrastive use in self-plagiarism for language models. In Lun-Wei Ku, Andre Martins, and Vivek Srikumar, editors, *Findings of the Association for Computational Linguistics: ACL 2024*, pages 14432–14447, Bangkok, Thailand, August 2024. Association for Computational Linguistics. Accessed March 2, 2026. URL: https://aclanthology.org/2024.findings-acl.858/, doi:10.18653/v1/2024.findings-acl.858.

[12] Pengfei Liu, Weizhe Yuan, Jinlan Fu, Zhengbao Jiang, Hiroaki Hayashi, and Graham Neubig. Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing. *ACM Computing Surveys*, 55(9):195:1–195:35, 2023. Accessed March 2, 2026. doi:10.1145/3560815.

[13] National Academies of Sciences, Engineering, and Medicine. *Fostering Integrity in Research*. The National Academies Press, Washington, DC, 2017. Accessed March 2, 2026. doi:10.17226/21896.

[14] Nature Editorial. Tools such as ChatGPT threaten transparent science; here are our ground rules for their use. *Nature*, 613:612, January 2023. Published January 24, 2023; accessed March 2, 2026. doi:10.1038/d41586-023-00191-1.

[15] OpenAI. A Student's Guide to Writing with ChatGPT. Website: https://chatgpt.com/use-cases/student-writing-guide, 2026. Accessed March 2, 2026. URL: https://chatgpt.com/use-cases/student-writing-guide.

[16] OpenAI Help Center. How can ChatGPT be used for assessment and feedback? Website: https://help.openai.com/en/articles/8313397-how-can-chatgpt-be-used-for-assessment-and-feedback, 2026. Accessed March 2, 2026. URL: https://help.openai.com/en/articles/8313397-how-can-chatgpt-be-used-for-assessment-and-feedback.

[17] teorth/erdosproblems Wiki contributors. AI contributions to Erdős problems. GitHub Wiki: https://github.com/teorth/erdosproblems/wiki/AI-contributions-to-Erd2026. Last edited March 1, 2026; accessed March 2, 2026. URL: https://github.com/teorth/erdosproblems/wiki/AI-contributions-to-Erd%C5%91s-problems.

[18] H. Holden Thorp. ChatGPT is fun, but not an author. *Science*, 379(6630):313, January 2023. Accessed March 2, 2026. doi:10.1126/science.adg7879.

[19] Trieu H. Trinh, Yuhuai Wu, Quoc V. Le, He He, and Thang Luong. Solving olympiad geometry without human demonstrations. *Nature*, 625(7995):476–482, January 2024. Published online January 17, 2024; accessed March 2, 2026. doi:10.1038/s41586-023-06747-5.

[20] UNESCO. Guidance for generative AI in education and research. Website: https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research, 2026. Accessed March 2, 2026. URL: https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research.

[21] World Association of Medical Editors. Recommendations on Publication Ethics Policies for Medical Journals. Website: https://www.wame.org/recommendations-on-publication-ethics-policies-for-medical-journals, n.d. See the Plagiarism section; accessed March 2, 2026. URL: https://www.wame.org/recommendations-on-publication-ethics-policies-for-medical-journals.

[22] Kaiyu Yang, Aidan M. Swope, Alex Gu, Rahul Chalamala, Peiyang Song, Shixing Yu, Saad Godil, Ryan Prenger, and Anima Anandkumar. LeanDojo: Theorem proving with retrieval-augmented language models. *CoRR*, abs/2306.15626, June 2023. arXiv preprint; accessed March 2, 2026. doi:10.48550/arXiv.2306.15626.