Professional norms in mathematical research in the AI age

Pavel Etingof (MIT)[1]

In [1] we argued that as AI surpasses humans in mathematical research, the ownership culture, which has long served as a basis of the reward system in mathematics at all levels, will eventually decline. However, this will not happen overnight, and it is not clear how long this process will take. The dampening factors are the inertia of the system (inherent in any long-established social structures), economic issues, and the possible safety concerns that may slow down the release of progressively stronger AI models to the public. We are in a transition period when the ownership culture is still present (even if in increasingly watered-down form), and reckless actions involving AI (such as racing, flag-planting, or scooping) may still devalue people's long-time work and harm their careers. This transition period is also marked with inequality of level of access to AI models by mathematicians: some have access at best to basic models, some to the most powerful public models, and some to even more powerful internal models. Also a number of mathematicians are not willing to use AI for mathematical research at all. Thus this period calls for updating the professional norms to include situations involving AI. These norms will likely have to be updated further as the situation develops, and perhaps down the road they will matter less, but at the moment adopting such norms seems important.

Here I would like to propose a set of norms (in line with the recommendations of [2]) that I believe are appropriate for established mathematicians, and that I intend to follow myself, at least in the near term.[2] I emphasize that these are just some suggested good practices, not a code of conduct that everyone is expected to follow. Each mathematician should feel free to adopt them or not, according to their own judgment.

1. Choice of research questions to explore using AI. When choosing the problem for tackling with AI with a plan to publicize the results, the researcher (henceforth R) should

(a) only choose problems in which they have a genuine mathematical interest, not merely well-known open problems;

(b) make a good faith effort to find out if other people or groups are already working on the same problem. If yes, R should not start working on this problem competing with them. On the other hand, R can offer collaboration if they can bring something essential into the project. For example, if R has better AI access, then they can offer coordination, but should not take credit unless they have significant human mathematical input. In this case R should try to play an auxiliary role and generally avoid co-authoring the work if the others are early-career mathematicians. This is especially important if R has access to an internal AI model which is not publicly available. The bar of signing work involving essential use of such a model (in terms of intellectual contribution of the signer) should be very high.

More generally, when someone has devoted substantial time and effort to a problem, R should take care not to mess up their efforts, especially if R has not made a comparable investment in the subject. This was, in fact, a widely accepted norm in the pre-AI era, but AI has made it much easier to violate.

2. Human-authored papers. Some technical work and writing in a human-authored paper may be done by AI, as long as

(a) each author has made significant intellectual contribution (ideas, techniques, or exposition);

(b) the paper is brought into a human-readable form and checked by AI agents and the authors, so that every detail in the paper is understood by at least one of the authors.

(c) the AI use is clearly acknowledged in detail, explaining which intellectual contributions are due to AI and which are due to the authors.

As in AI-free papers, the authors take full responsibility for correctness of the paper and completeness of references. The contributions of individual authors need not be discussed, they just need to be separated from AI contributions. The authors may include the AI model as a co-author if they wish, with the understanding that under the current rules in this case the ArXiv and most journals will not accept it.

[1]Disclaimer: I have been part-time consulting for Anthropic since July 14, 2026. However, this text represents my own views consistent with my recommendations in `https://math.mit.edu/~etingof/aiuse.pdf` (May 2026). My views on the subject have not changed since then.

[2]I am grateful to Henry Cohn, Slava Gerovitch, and Andrew Sutherland for useful comments.

3. AI-authored and author-free notes. If one of the conditions (a),(b) in the previous subsection is not satisfied, it may be reasonable to publicize the paper on a personal webpage or, better, a site like `https://hexagonmath.org/` (which is more visible and has version history), as a note without authors or with AI model being the author. In this case the human mathematician submitting the paper serves as a contributor, and should explain AI use in detail and confirm a good faith effort to check the correctness of the paper and the completeness of the references (at least agentically). An agentic reference check should be done the day of posting for both kinds of papers, to avoid the kind of embarrassment that occurred with the October 6 Open AI release. Sometimes it even makes sense to publish the whole sequence of prompts; this may be useful for other researchers. The contributor should also indicate the degree of human understanding of the paper (e.g., in `https://hexagonmath.org/`, there are three levels: none, some, or full, which may be upgraded at a later date). The contributor should not get academic credit for such notes beyond it being a service to the community. The contributor and other researchers should be free to build on such notes with a suitable reference. This mode of dissemination can be used, for instance, for less important results which are still good to know for the community, or ones for which the contributor has currently no bandwidth to meet the criteria of a human-authored paper. Such notes will be read primarily by AI agents or people who need a specific result in their work.

Provided that an AI-generated note clearly describes its provenance, the extent to which humans understand its content, and says who, if anyone, takes responsibility for it, there seems to be little difference between naming AI as its author and presenting the note as author-free. I think this should be a matter of choice.

4. Counterexamples. If R discovers a counterexample to a conjecture in the literature (with or without AI), but this counterexample would not be a significant standalone mathematical result, they can post the example on `https://hexagonmath.org/` or personal webpage in one of the two ways described above. However, if the counterexample itself is a major result and there may be junior researchers who have spent a lot of effort looking for it, R should make a good faith effort to find such researchers and possibly offer coordination. In all cases, R should first get in touch with the author of the conjecture, if possible.

5. Interesting results. If collaboration with AI has led R to interesting (or even major) results on some problem, with most work done by AI, R should think about other researchers or groups who may be working on this problem and would be interested in the results, and if they exist, consider sending them AI output and offer coordination. R should not take credit for this progress unless they make a genuine intellectual contribution, beyond just prompting AI to work on the problem. Rushing, flag-planting or intentional scooping other mathematicians are inappropriate, and getting into priority disputes is even more pointless and ill-advised than it was in the past.

6. Communications. R should not deliberately put into AI any private communications from other mathematicians without their consent, especially with the intention of producing and publicizing research results. Similar restraint is strongly recommended for abstracts of talks and comments made in talks or informal conversations.

7. Mistakes. If in the course of research R discovers a mistake in someone's paper (with or without AI), they should first contact the authors if possible and give them time to acknowledge it and fix it. It is nice to ask AI to propose corrections and send them to the authors; R should not take credit for any of this if most of it is done by AI. If this is not feasible (e.g. the authors are no longer doing mathematics or are unresponsive), it may make sense to put a note on `https://hexagonmath.org/` according to the above rules which points out the mistake and suggests fixes if possible.

[1] P. Etingof, S. Gerovitch, Mathematics of the future, October 5, 2026, `https://math.mit.edu/~etingof/futuremath.pdf`

[2] Report of Summit on Math Research in the Age of AI, October 22-23, 2026, `https://cmsa.fas.harvard.edu/aimathresearch_summit/`