Qeios

Peer Review

Review of: "The Role of Artificial Intelligence in the Lifecycle of Scientific Manuscripts: Authoring, Reviewing, and Editorial Selection"

Nicholas Baker[1]

1. University of Windsor, Canada

Thank you for the opportunity to participate in this open review process. By way of acknowledging other reviewers, my own process has been to independently review the article, then check themes in others' reviews so that I can focus on providing mostly new observations and feedback. Having said that, I do feel there are also some points that I would like to reinforce that others have made already.

Strengths

This is an interesting paper and addresses two issues that have been increasingly prominent in recent years – the for-profit publishing system, and the potential impact of generative AI on that system. I would like to reinforce some of the points that the paper makes that others have also identified:

1. Commercial, for-profit publishing is a very strange model that is not optimised for the creators and end users of the knowledge that is shared through these mechanisms. The author does a good job of highlighting this issue and attempts to offer ways to adapt what we have now to something fairer and a better fit for purpose for the academic community.
2. Most of the references are very recent, which is important with the topic of AI (though as I will note below, even with this, many of the arguments about AI's capabilities are not accurate or simply out of date because this field is evolving so rapidly).
3. The tables provide useful quick overviews/syntheses of the evidence in a non-systematic sense, though with the caveat that much of the AI content is out of date. I think the addition of a table that summarised major publisher/journal policies on the use of AI and their compensation model + APC fees would potentially add a lot to the article.

doi.org/10.32388/P2GKTS

4. The hybrid model proposed is, I believe, a useful way forward for academic publishing, though I don't think it necessarily needs to include additional compensation for reviewers, and many journals have already adopted this model.

Additional feedback not already provided

As an overall comment on articles attempting to address the challenges and capability gaps in AI, I am seeing a pattern in much of what I'm asked to review that seems reminiscent of those written about online and digital learning – there is a level of scrutiny being applied to AI that is not also applied to the human-only counterpart (i.e., similar to the in-person vs online dichotomy where online is heavily scrutinised and in-person is never questioned seriously as the gold standard for learning). For example, there is much discussion about hallucinations, lying, inaccuracies, lack of creativity, poor writing abilities, variability in answers, bias, application of dominant worldviews, etc., that are all characteristics of human society and academia, yet we don't seem willing to remove humans from processes that lead to or are impacted by all those issues. In my view, there needs to be much more tolerance for exploration and mitigation, rather than rejection of AI simply because it displays many of the same characteristics as its human creators, and a much more critical – really critical rather than uninformed criticism – discussion of AI in academic work if we are to navigate the changes it is precipitating whether we like it or not. I think we also need to acknowledge that academics have for decades been outsourcing the writing of journal articles to people not acknowledged as authors, including those who do so for profit and have not made any serious attempt to eliminate the practice, so this is no different from completely outsourcing work to AI in my view.

Some additional comments:

1. Much of the academic publishing world is indeed driven by profit, but there are alternative models such as in-house journals, fully open journals, and society-led journals that already exist and are quite successful. The author gives a nod to at least the last of these, but there is no acknowledgement that all of those alternative models also have costs to them. There is much more unpaid labour that goes into running your own fully open journal with no APCs to support it, and there is an additional burden on the society, the university, or fundraising to support the actual costs of the publication, e.g., hosting, some professional editing support, development costs, etc. I would also suggest it is worth drawing the distinction clearly between fully open non-profit journals that

survive on APCs, and for-profit journals that also charge APCs (and much larger APCs for publishing openly). This is not a problem of openness; it is a problem of how that is resourced.

2. While I agree that many of the APCs (and especially if you want to publish openly) are outrageous, there is no analysis offered here of where those costs go and what the professionally/commercially operated journals offer that other models don't. In short, I'd like to see a cost-benefit analysis, even a basic one, that supports the argument for an alternative model. As others have noted, as part of that there should be some justification for the suggested compensation to reviewers, and acknowledgement that many journals already offer incentives, particularly through credits towards APCs, public recognition, or access to other published works for their reviewers.
3. I also think it is important to acknowledge that the big expensive journals that the author attacks and is arguing against exist in part because we in academia built the model they thrive on. We value research and publication above all other academic activities, leading to ever more pressure to publish, leading to ever more articles being written, more metrics being invented, and so on, and the publishers to some extent are just giving us what we have shown we actually value. Academia requires a significant culture shift if we are to actually address the issues raised by the author.
4. In section 2.2, the author suggests that the 'profit-driven atmosphere is the real reason that AI is being introduced to the peer-review process.' There is no evidence of that, and in fact, I would argue that it is quite the opposite. AI is typically being applied to workflows that humans would rather not do themselves, that they don't value enough to do themselves, or which they do not have enough time to do themselves. There is no indication that journals are making widespread moves to introduce AI review processes (most still only allow its use in preparing manuscripts), and it would actually add additional cost to the work of the journal because they would need to provide the AI service in order to make it equitable and trustworthy, which would require investment. I'm not suggesting there is not a potential real benefit to developing some AI assistant that checks first drafts prior to submission to make sure that the references are all real, that gives feedback on or fixes formatting errors, that formats references, that offers suggestions on grammar/spelling/flow, etc., as I think there will be those types of tools available in the short term (some universities have already built their own), but I don't see any evidence that it goes beyond that, at least in the short term.
5. What the author does not acknowledge is the fact that many stressed, overworked, potentially lazy, potentially oblivious academics are already using AI to write their reviews, completely independent of the publishers and usually in breach of the publishers' rules. That is a much larger threat to

academic publishing than if the publishers build their own trusted systems. Additionally, it seems likely in the AI-enabled environment that if relatively large incentives to review were introduced as the author suggests, it would lead to perverse outcomes whereby some would take advantage of AI to accept many review requests, simply run the paper through AI, and collect the cash.

6. As a couple of the other reviewers have noted, the section on the issues associated with AI is considerably inaccurate and does not align with either the capabilities of current models or the actual critical use of AI in scientific publication. I see a few reasons for that:
    a. First, even though the references are quite recent, at the pace that capabilities are improving, a paper from 2023, or even 2025 if it has taken a year to get published, is so far out of date as to be only useful as a point in time indicator. It is possible to speak to the issues at that time, as long as it is acknowledged that at the current time point when this article is submitted, many of those issues no longer exist (or perhaps new ones have appeared). It is important to be precise when talking about AI issues, and this lack of precision is a constant issue in the current literature.
    b. All of the evidence cited for the issues with AI refers only to generalist tools, and almost all of it to ChatGPT. Anyone who is critically using AI in scientific publishing who is not using a specialised tool of the types suggested by one of the other reviewers is just lacking the necessary AI literacy to choose an appropriate system and/or use it well. This will rapidly change as AI literacy develops in the scientific community.
    c. The author does make a very good point about cultural and geographic differences in AI accuracy when using some of the mainstream Western-developed models, which is worth exploring in a little more detail, but even there, it must be acknowledged that this is one of the reasons that sovereign models are being developed globally, and that even the big Western models have both improved and are working very hard to increase knowledge capabilities in non-dominant languages and cultures as a standard part of their development cycles now.
    d. There is some reference to specialised areas that are not well-served by current AI models, but none of those are actually indicated or have details provided. It would be very helpful to include some examples, especially if they are current and refer to the current state of the art and not models from years ago.
    e. In section 3.2, the author notes that the entire responsibility for [detecting, preventing] fraud is on authors, reviewers, and editors. I would agree, but also argue that this has always been the case and should continue to be so. It is the responsibility of the human author to ensure that

anything they submit for review is accurate and does not contain AI-generated errors. Current publishing guidelines make it clear that this is the case. This same section also has the sentence “Although the adoption of RAG tools offers a partial shield, improving accuracy, it cannot replace true comprehension, leaving the fundamental risk of AI-generated misinformation unresolved.” First, I would suggest that RAG and inaccuracy have nothing to do with ‘true comprehension’ – those concepts are not equivalent. RAG is an approach we can use (one of many) that improves accuracy by bounding the information the model checks against when it is reasoning against a problem. It still does not absolve the author of the responsibility of checking their work (or the work of their AI assistant), and this is only really a problem if the author’s intention is to ask an AI assistant to do all the work for them and remove themselves from the cognitive work altogether – that I would argue is a much bigger issue than whether the tool they choose makes stuff up or not!

7. In section 4.1, point c, LLMs are actually not very good at plagiarism detection (other deterministic algorithmic methods are better at this work and have been used for decades for that purpose).
8. Section 4.2: there is no evidence offered for the things that are claimed as irreplaceable human roles, and in fact, for each of these, there is evidence to the contrary. For example, there is a growing body of evidence that shows that current frontier models are more creative than all but the very highest-performing creative humans on objective measures of creativity that have been validated; there is evidence that AI applies ethical principles and values more consistently and rigorously than almost all humans across many different domains; there are whole tools that are designed to help researchers understand the novelty and significance of their ideas, e.g., what does the literature say about a topic, are there interdisciplinary or edge cases that can contribute to the thinking, has this approach been taken before, does a patent exist for anything similar, etc. All of that exists now, and while I agree that these are things we would like to believe are uniquely human or that humans will always be better at, the evidence suggests otherwise (which is absolutely contributing to the existential angst so many of us are currently experiencing in academia).
9. As others have noted, there are some conceptual issues with figure 1, particularly with how the arrows appear to represent the publishing process. I wondered immediately (confirmed in the acknowledgements) if the image was AI-generated, and if so, that should be acknowledged in the text itself for transparency. It also has two titles/captions (the one in the image itself should be removed), and the text in the bottom right of the image is actually part of the prompt, so needs to be

removed. In the paragraph just before the image, AI is described as a 'precise' tool, whereas it has been suggested throughout that it is the complete opposite, so that needs to be adjusted.

10. Table 2 would benefit from references to the literature, especially as there are several suggestions (as indicated above) that are inaccurate.
11. As noted above as well, for section 5, I don't think the options are only to pay the reviewer or not. Why not also pay the author, for example? Why not reward in other ways?
12. Section 5.2: The argument of getting publications out quickly is definitely a trend we are seeing with many journals expediting processes and the rapid rise of pre-print servers. That was happening before AI, though. It may be worth noting the impact of predatory journals on that process as well – these have pushed legitimate publishers to speed up processes (and they also demand large APCs for publication).
13. Section 5.3 makes the argument that replacing human reviewers with AI would save costs, but there is no indication of how that would happen (especially if reviewers are already unpaid). This section also refers to opaque algorithmic decision processes, but that does not have to be the case – it can, in fact, be highly transparent (and I would argue for that as a necessary condition if AI is to be used in the review process). This section also repeats the unsubstantiated claim that AI is not capable of identifying novelty or significance, and adds that innovative work from unknown labs or on emerging topics would be systematically rejected – there is absolutely no evidence that this is the case, and additionally, no evidence that it is not happening now with human reviewers. I would suggest that this whole section is arguing against practices that are already happening, and somehow assumes that no one (human) would be responsible for decision-making, which at present is also highly unlikely.
14. Section 6 opens with a sentence about editors using predictive analytics to select manuscripts, which I had not heard of as a practice and was interested to learn about. The reference to which this is ascribed, though, is one I have read before (and checked again) and has no relationship to the claim, so it appears the incorrect reference was applied. If this practice is happening, which is disturbing, I would suggest that it was already happening, just without the aid of algorithmic data – editors inherently decide what is and is not likely to be a good fit for their journal's audience, so they are already gatekeeping based on their experience of what has been published in the journal in the past, their expertise, and gut feeling.
15. Section 7.1 says that publishers have established AI policies that universally prohibit AI authorship – that's not quite true. AI can be used in many publications in a role that is very similar to traditional

authorship; it just can't be acknowledged as an author for most, usually with the caveat that the AI cannot take responsibility for the work whereas human authors can. These author guidelines have softened somewhat in the last year or so and become more nuanced. It may be worth including another table that provides some examples of different journal guidelines across different disciplines (or publishers) as that would help the reader understand the variance that exists now and what is generally thought of as acceptable practice.

16. Section 7.3 provides a list of things that researchers should be trained to do – I would argue that none of them are actually AI literacy but rather represent academic integrity values and/or information/data literacy. There are, however, many ethical and AI literacy frameworks that already exist for AI in research and higher education that could be referred to here.

Declarations

Potential competing interests: No potential competing interests to declare.