Comment l'IA générative transforme les pratiques de recherche : nouveaux enjeux d'intégrité scientifique

Ofis

Colloque 15 avril 2025

IPGP - Amphithéâtre
1, rue Jussieu - 75005 Paris

AI tools in peer reviewing: challenges and needs

Vasiliki Mollaki, PhD, MSc
National Commission for Bioethics & Technoethics, Greece
National Centre for Scientific Research "Demokritos", Greece

vmollaki@iit.demokritos.gr

15 April 2025

Importance of the Peer Review Process (PRP)

- Scholarly communication
- Sets academic research to the judgment of expert peers to critically assess the:
  - Novelty
  - Quality
  - Impact of research
- Quality control of scientific research
- Ethicality and integrity of scientific research

Guston, 2007; Rennie 2003; Ware, 2008

Advantages of using AI tools in the PRP

- Enable **Editors** to:
  - Accelerate manuscript screening
  - Check if manuscripts fit with journal scope
  - Detect plagiarism
  - Check formatting
  - Assess compliance with journal policies
  - Identify suitable reviewers
  - Summarize individual review reports
  - Write final decision letters

- Enable **Reviewers** to:
  - Generate more concise, informative and well-written reviews (particularly for non-native speakers)
  - Save time and effort from correcting grammar and spelling mistakes

Can Generative AI be an author?

- Increasing use of AI-assisted technologies (LLMs) in the manuscript preparation phase (Dergaa *et al*., 2023; Lee *et al*., 2023)
- This enhances and speeds up academic writing
- But it also raises ethics issues on (Dergaa *et al*., 2023):
  - authorship,
  - authenticity,
  - credibility and accountability of academic work,
  - with subsequent legal implications on copyright

Policies of large publishers* on authorship

*by journal count Nishikawa-Pacher, 2022

Publisher	Policy on the use of AI-based tools for authors
Springer	LLMs, such as ChatGPT, do not currently satisfy the authorship criteria Use of an LLM should be properly documented in the manuscript
Taylor & Francis	**Committee On Publication Ethics (COPE) guidelines** AI tools, such as ChatGPT or LLMs, cannot be listed as an author of a paper, disclosure
Elsevier	Authors allowed to use generative AI and AI-assisted technologies in the writing process before submission, but only to improve the language and readability of their paper and with the appropriate disclosure
Wiley	COPE guidelines The final decision about whether use of an AI tool is appropriate or permissible lies with the journal's editor or other party responsible for the publication's editorial policy
SAGE	COPE and **World Association of Medical Editors (WAME) guidelines** AI bots, such as ChatGPT, should not be listed as an author. Authors who use AI tools must acknowledge this in the manuscript. Editors and reviewers should evaluate the appropriateness of the use of LLMs
OMICS	None
DeGruyter	None
Oxford University Press	Neither symbolic figures nor natural language processing tools driven by AI, such as ChatGPT, qualify as authors. The use of AI tools must be disclosed in the manuscript and the cover letter
Inder Science	COPE guidelines AI tools (e.g. ChatGPT) cannot be listed as authors. Authors are fully responsible for the content of their article, even those parts produced by any AI tool, and are thus liable for any inaccuracies or breach of publication ethics. The use of AI tools must be acknowledged in the manuscript
Brill	COPE guidelines Authors may use AI and LLMs in the writing and preparation of their manuscripts when doing so with transparency, maintaining full responsibility and accountability for their research

Mollaki 2024

Can AI be a reviewer?

- What if peer reviewers use AI tools for their peer review reports?
- What would be an **appropriate policy** for editors to follow in the case that authors expressed their concerns, or editors themselves have suspicions that the peer review report is a result of, or it includes parts produced by, generative AI tools?
- Are these sufficiently described in journal publishing policies?

Policies of large publishers* on peer reviewing

*by journal count Nishikawa-Pacher, 2022

Publisher	Policy on the use of AI-based tools	
	for editors	for reviewers
Springer	None	None
Taylor & Francis	Files, images or information from unpublished manuscripts must not be uploaded into databases or tools that **do not guarantee confidentiality**, are accessible by the public and/or may store or use this information for their own purposes (for example, generative AI tools like ChatGPT)	**Reviewers must not use AI tools** to generate manuscript review reports, including LLM-based tools like ChatGPT
Elsevier	None	Reviewers should not upload a submitted manuscript or any part of it into a generative AI tool as this may violate the authors' **confidentiality and proprietary rights** and, where the paper contains personally identifiable information, may breach data privacy rights
Wiley	None	None
SAGE	None	None
OMICS	None	None
DeGruyter	Refer to journals' policy	None
Oxford University Press	None	None
Inder Science	None	None
Brill	Editors same as authors	None

Mollaki 2024

Risks of using AI tools for manuscript reviewing

Donker 2023 on ChatGPT

- Good summaries
- No specific improvements on the study
- No critical content on the study
- Specific-looking general comments with no bearing on the text
- Specific but unrelated comments to the study content (reasons for rejection)
- **Fabricated references** (non-existent but authored by real persons working on similar topics)
- Could be mistaken for an actual review report by persons that have not fully read the manuscript
- Authors should be prepared to challenge reviewer comments that seem unrelated and non-specific

Checco *et al*. 2021 on an AI tool

- AI tool developed to assess if AI can approximate human decisions in PRP
- It was often able to successfully predict the peer review outcome reached as a result of human reviewers' recommendations
- The "first-impression bias" found in reviewers can be present in a training dataset and thus, the trained **AI model can propagate such biases**

Hosseini & Horbach 2023

- AI tools could allow **fake** peer reviewers to create more unique and well-written reviews

Risks of using AI tools for manuscript reviewing

- Need a consistent, **end-to-end policy** on the ethics and integrity of AI tools in publishing, to avoid the risk of compromising the integrity of the PRP and of undermining the credibility of academic publication

Garcia, 2023; Ling and Yan, 2023

- Risk of editors and reviewers over-relying on AI tools to assess the originality, quality and impact of a manuscript, overlooking their own experience and expert judgment
- Dependency on AI tools -even to a certain extent- to manage the PRP can jeopardize the way editors and reviewers exercise their **autonomy**
- Consequently, authors may **distrust** the PRP due to a lack of transparency on the rationale of decision-making
- AI tools used in the PRP may **undermine the integrity and the purpose** of the process, challenge academic communication and even trust in science at large

Mollaki 2024

Are policies enough?

Policies for use of AI tools by authors are restrictive

Policies for use of AI tools by peer reviewers are becoming (or will become) restrictive too

But…are policies enough? What steps need to be taken to ensure that policies are followed?

Beyond policies and call for action

Authors suspect that a reviewer has used AI tools to produce a review report, what should they do?

Alert the editors, who will raise such matters to the journal's Ethics Committee (assuming one exists)

How can we prove whether allegations are true? How can we fill in the gap of **evidence**?

In the absence of evidence, how could such allegations be investigated by Ethics Committees?

Policies & procedures must be **transparent**, detailed and **solid**, facilitating decision-making when a reviewer is found to have used AI-tools

Beyond policies and call for action

WAME guidelines for the use of AI tools by authors recommend that:

"[. . .] **editors need appropriate tools** to help them detect content generated or altered by AI for the good of science and the public, and to help ensure the integrity of healthcare information and reducing the risk of adverse health outcomes" (Zielinski *et al*., 2023)

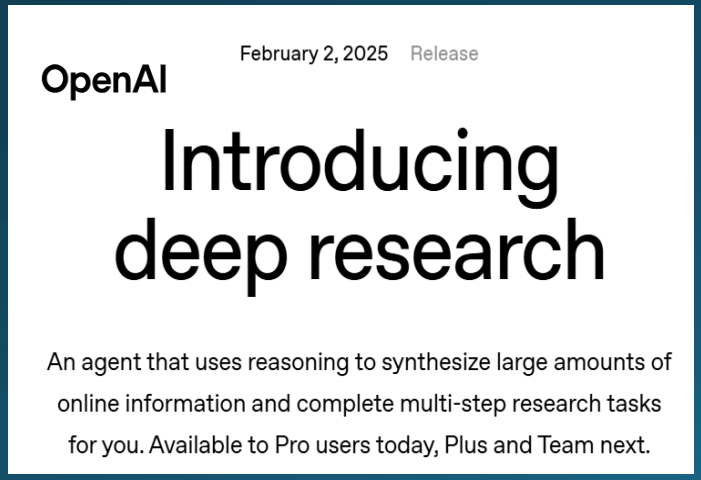
- Should we use similar **tools to detect** AI-generated/altered peer review reports?
- Will this undermine the altruism on which peer review is based on? (reviewers voluntarily spend time and effort to help authors improve their papers, Mulligan & Raphael, 2010)
- Checking for AI content in review reports **challenges the trust between reviewers and editors**, and the reviewers' willingness to continue offering expertise in the PRP

Beyond policies and call for action

- If the use of AI tools in review reports cannot be proven, e.g. through tools that detect content generated/altered by AI, it will be **detrimental for ethics in publishing**
- Adding on existing criticism on the PRP: long delays, efficacy of the process, susceptibility to bias, inability to detect fraudulent research/ misconduct (Castelo-Branco, 2023; Manchikanti *et al*., 2015; Tennant *et al*., 2017)
- Reviewers found not to comply with publishing policies must be **flagged and excluded** from the PRP, for all journals under the same publisher
- Publishing policies are essential, but they need to **enable certain solid steps** towards protecting peer review ethics. Reluctance to do so can:
  - put the journal's and/or the **publisher's reputation at risk**
  - compromise **the integrity and ethics of the peer review system**

Why is it important to have policies & procedures that are transparent and solid?

- Everybody wants to publish, nobody has time to review !
- AI tools are regularly optimized, thus, they become better
- Consequently, papers written by AI tools become better (e.g. ChatGPT deep research)



Papers written by AI tools will be reviewed by … AI tools !

Loss of PRP's scope "to critically assess the novelty, quality, and impact of research by peer experts"

Thank you

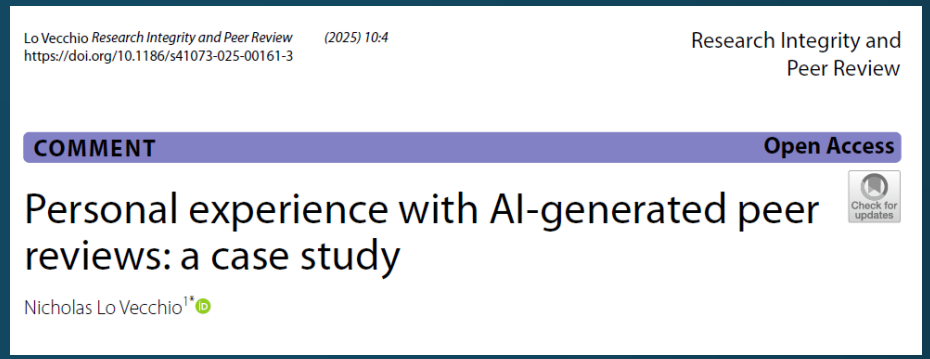
Special collection: Research Integrity and Research Misconduct

RE

Death of a reviewer or death of peer review integrity? the challenges of using AI tools in peer reviewing and the need to go beyond publishing policies

Research Ethics
2024, Vol. 20(2) 239–250
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/17470161231224552
journals.sagepub.com/home/rea

Sage



Lo Vecchio *Research Integrity and Peer Review* (2025) 10:4
https://doi.org/10.1186/s41073-025-00161-3

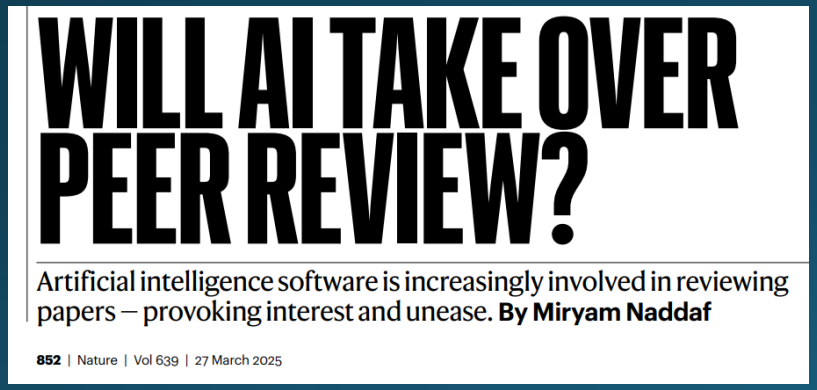
Research Integrity and Peer Review

COMMENT Open Access

Check for updates

Personal experience with AI-generated peer reviews: a case study

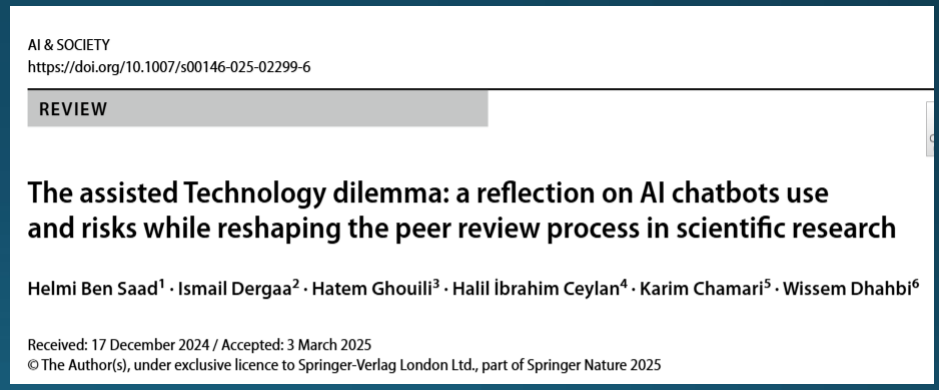
Nicholas Lo Vecchio[1*]



WILL AI TAKE OVER PEER REVIEW?

Artificial intelligence software is increasingly involved in reviewing papers — provoking interest and unease. **By Miryam Naddaf**

852 | Nature | Vol 639 | 27 March 2025



AI & SOCIETY
https://doi.org/10.1007/s00146-025-02299-6

REVIEW

The assisted Technology dilemma: a reflection on AI chatbots use and risks while reshaping the peer review process in scientific research

Helmi Ben Saad[1] · Ismail Dergaa[2] · Hatem Ghouili[3] · Halil İbrahim Ceylan[4] · Karim Chamari[5] · Wissem Dhahbi[6]

Received: 17 December 2024 / Accepted: 3 March 2025
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2025

vmollaki@iit.demokritos.gr