

Correspondence

Artificial Intelligence Authorship— Conscious Intent, Moral Agency, False Accountability, and the Value of Authorship Credit

Bor Luen Tang 

Department of Biochemistry, Yong Loo Lin School of Medicine, National University
Health System, National University of Singapore, 8 Medical Dr, Singapore

bchtbl@nus.edu.sg

orcid.org/0000-0002-1925-636X

Received: 3 Sep 2024
Accepted: 30 Sep 2024
Published: 4 Nov 2024

Declaration of Interests

The author has no conflicts of interest to declare.

Funding

The author declares that this article received no financial support.



This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0).

Citation

Tang BL. Artificial intelligence authorship—conscious intent, moral agency, false accountability, and the value of authorship credit. *Eur Sci Ed.* 2024;50:e136216.

<https://doi.org/10.3897/ese.2024.e136216>

A recent viewpoint article¹ argued that artificial intelligence (AI)/large language models (LLMs) proposed “AI-authorship,” with modifications made to the ICMJE guidelines. Here, I offer a few counter perspectives and caveats against the possible implementation of AI-authorship.

Firstly, the barring of LLMs and AI chatbots as authors or coauthors on academic papers is not primarily because of a fear of the latter being intellectual threats, but rather because AI tools cannot act autonomously with conscious intents. An AI LLM will only produce relevant content in response to precise human prompts, and are non-starters as far as autonomy in making an intellectual contribution is concerned. Furthermore, although LLMs could appear to comprehend and generate human language text, they do so in a way that is apparently different from humans. LLMs work² based on conditional probability—the machines learn a high-dimensional probability distribution over sequences of words based on their training datasets, and then compute the most statistically probable word that should appear next in a sentence. While humans often encounter situations in which they find it difficult to express an idea in words, LLMs could always have the words even without an idea. Therefore, as language tools, LLM and AI chatbots are very proficient, but they cannot count as intellects due to the lack of autonomous agency.

Secondly, LLM and AI chatbots also lack moral agency. A moral agent is an entity that is expected to meet the demands of morality,³ and whose actions are thus eligible for moral consideration. While there has been a great deal of discussion and promises on the development and implementation of “responsible” or “trustworthy” AI, Conrad and Nagel have cautioned that we should not inappropriately attribute trustworthiness or responsibility to AI technologies.⁴ As AI technologies, some of which have more acute moral issues (autonomous vehicles,

facial recognition, healthcare decisions, etc.) that require urgent resolution, LLMs and AI chatbots used for academic writing might present comparatively milder issues. However, these are likewise not moral agents. Their actions are dependent on human prompts, and whatever content they generate will only be submitted and published by humans. Without the agency of autonomy, there cannot be any accessible moral responsibility.

Following from the above, it is therefore meaningless to talk about accountability, for after all, it is only the human wills and actions that shall bring about an AI-written/writing-assisted paper, with all its glory as well as biases, inaccuracies, and errors. In other words, only human authors, being moral agents, could be responsible and accountable for the content of a paper. Not being moral agents and thus not expected to meet the demands of morality, accountability carries no meaning for AI tools, to begin with. As such, any changes to the ICMJE authorship clauses⁵ just to accommodate AI-authorship would be a fallacious move that could jeopardize the true value of authorship credit.

Perhaps of even greater importance is how AI-authorship might affect human authorship. If the use of LLMs and AI chatbots becomes so widespread that 80-90% of papers are produced with the aid of AI and have an (or more) AI co-author, would that not cheapen if not displace human authorship? Consider the diligent graduate student who works long hours contributing to the lab or group paper at the expense of her own PhD project, only to find her name behind that of a chatbot, or the hard-working technician who has worked night and weekends to process samples while they are fresh, only to find her name in the “Acknowledgement” section while GPT-4o is an author because the latter wrote the draft and drew the figures. What is the point if most, if not all papers, have

almost by default an AI author? Would that be appreciated by AI chatbots (or their human manufacturers), and to whom are we extending equality to, if not a moral agent?

There might be a tendency to reward AI chatbots with authorship because of a sense of collegiality generated by the unusually intellectually stimulating interactions (to the human) during their use. However, beyond emotional attachment, is there a clear difference in terms of actual intellectual contribution between an AI chatbot and other critically contributing high-end AI-based research tools, such as AlphaFold2.⁶

References

1. Teixeira da Silva JA. Proposing authorship for artificial intelligence and large language models. *ESE*. 2024;50:e123908. [CrossRef]
2. Ramponi M. How ChatGPT actually works. Available at: <https://www.assemblyai.com/blog/how-chat-gpt-actually-works>. Accessed August 2024.
3. Vinit H. Moral agents. Available at: <https://www.rep.routledge.com/articles/thematic/moral-agents/v-1>. Accessed September 2024.
4. Conradie NH, Nagel SK. No agent in the machine: being trustworthy and responsible about AI. *Philos Technol*. 2024;37(2):72. [CrossRef]
5. International Committee of Medical Journal Editors. 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 September 2024.
6. Yang Z, Zeng X, Zhao Y, Chen R. AlphaFold2 and its applications in the fields of biology and medicine. *Signal Transduct Target Ther*. 2023;8(1):115. [CrossRef]

ease / publications

ese / European Science Editing

European Science Editing is an official publication of EASE. It is an open access peer-reviewed journal that publishes original research, review and commentary on all aspects of scientific, scholarly editing and publishing.

<https://ese.arphahub.com/>
<https://www.ease.org.uk>
https://twitter.com/Eur_Sci_Ed
<https://www.linkedin.com/company/easeeditors/>



© 2024 the authors. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.