COPE

publicationethics.org

DISCUSSION DOCUMENT:

ARTIFICIAL INTELLIGENCE (AI) IN DECISION MAKING

DISCUSSION

VERSION 1: September 2021

ARTIFICIAL INTELLIGENCE (AI) IN DECISION MAKING

COPE discussion documents introduce issues and aim to stimulate discussion about a topic. They are often about complex issues and **COPE** welcomes comments which add to the ongoing debate.

COPE discussion documents should not be considered formal **COPE** policy. Discussion documents may be revised but remain a discussion document as the issues develop and if further comment from the community is sought, or they may be revised to become formal **COPE** policy (and published as **COPE** Guidelines).

We welcome further comments; please send your feedback to **Natalie Ridgeway, Executive Officer** cope_execofficer@publicationethics.org

Introduction

Artificial intelligence (AI) technology has advanced rapidly over the past several years and has started to introduce data driven solutions into the publication process of scholarly articles. Some of the steps in the publication process where AI tools are already being used as support include journal selection, identification of the subject matter in a paper, determining if the subject falls within the scope of a journal, reviewer suggestions for peer review, language quality assessment, text overlap and plagiarism detection, document formatting, as well as assessing the appropriateness of experimental design and statistical analyses. Opportunities for AI use in publishing are increasing at a high rate and bring with them new challenges, as well as ethical questions that must be considered. This document represents a first attempt to address these questions and share recommendations. It is an elaboration of an initial **COPE** Forum discussion topic on "Artificial intelligence (AI) in decision making".[1] The aim of this discussion document is to provide guidance and general recommendations on the current state of the issue.

Cite this as: COPE Council. COPE Discussion Document: Artificial intelligence (AI) in decision making — English. https://doi.org/10.24318/9kvAgrnJ

Automation versus AI

There is a clear distinction between automation and AI, and the two should not be used interchangeably. An understanding of what AI in publishing truly represents is key to this discussion.

Automation refers to rules based software, typically ensuring that a process moves forward, without human intervention, based on a set of predetermined explicit rules. The rules are defined by the developers of the automated system. An example of automation in peer review is an editorial management software system that automatically sends email reminders to authors who are late to resubmit their manuscript, or the system holding a manuscript from processing to the next stage of peer review if a specific file is missing.

AI is about engineering intelligent systems, machines, and software that can mimic human intelligence and behaviour. The aim is to augment (and even supersede) human intelligence, using large quantities of data to create algorithms, neural networks and graphs, and deep learning technology to achieve a level of so-called intelligence that is greater than what a human mind could achieve on its own in accuracy or scale. We include natural language processing (NLP) and machine learning (ML) as subsets of AI. NLP is how computers process and attempt to understand written or spoken text and perform tasks such as keyword extraction and topic classification; ML applies algorithms trained on datasets to identify patterns to make predictions, perform tasks, or take decisions without having been explicitly programmed to do so. An example of AI in peer review is systems providing lists of peer reviewers for a manuscript, using ML on large researchers' databases and NLP for text processing of articles, and sending invitations to review the manuscript to the most appropriate selection. AI ultimately provides valid recommendations that a human editor could have achieved, using a larger knowledge base and on a broader scale.

AI powered software can provide results to be used in an automated system, or to be assessed and decided on by an editor or author. For example, AI technology can predict which journal is the better match in scope for a manuscript that has been submitted to a publisher. The system can then automatically assign the manuscript to this journal, without a human in the loop, which would constitute an AI based decision. Another example is text overlap detection results, provided by software such as iThenticate,[2] being used to make an automatic decision to reject a manuscript for plagiarism, without human verification. These are examples of AI powered automation in decision making.

In summary, AI tools can be developed to provide guidance to humans, based on relevant data; AI could also support full automation of some processes and decision making without human intervention. In the ethical dilemmas section below, we raise the ethical implications associated with AI making unsupervised decisions in the publishing process.

Cite this as: COPE Council. COPE Discussion Document: Artificial intelligence (AI) in decision making — English. https://doi.org/10.24318/9kvAgrnJ

Version 1: September 2021.

Why use automation and AI in publishing?

Automation in publishing has been used for decades, to ensure manuscripts can be peer reviewed more swiftly, without the use of human intervention at every step of the process. The standard examples include systems sending reminders of various tasks to authors, reviewers, and editors. Separately, and more recently, AI has demonstrated its potential to solve problems that are not easy (or sometimes even possible) for a human mind to solve, or would require an unmanageable amount of time to complete. The AI and automation tools being developed have the power to help with the speed and accuracy of peer review. Software created to detect text overlap provides a level of assessment, by cross checking millions of documents, that a human brain could not achieve. The AI possibilities for pattern recognition make it possible to detect citation cartels, image manipulation,[3] salami slicing, and papermill characteristics,[4] all problematic non-ethical practices in publishing that are difficult to identify. The latest technology can assess, without the need for human input, whether a manuscript meets quality standards for language, format according to journal standards, presentation of figures, and use of citations.

In publishing, automation has been used mostly to move manuscripts forward through the peer review process with the least human intervention for processes that are not providing final decisions on manuscript publication. With AI, various attempts are being made to give decision making power to the artificial intelligence to make the final decision on articles for acceptance or rejection.[5] It can be argued that such AI tools might remove a level of personal bias that comes with interventions from human editors (eg, prejudice towards certain authors, or country specificity in invitations to reviewers). However, AI can have other inherent biases based on the data it was trained on, who developed it, and the software design itself.[6 7]

The overall aim of using AI assisted automation in publishing is to improve the quality of the automation, lessen the burden on human participants, and increase the speed of the review process, ultimately sharing validated and peer reviewed research outcomes more quickly, and with reduced demands on editors, reviewers, and authors.

Context and challenges

With the advancement of AI, questions arise about the ethics of when and how AI could, and should, be used for autonomous decision making. Several organisations and governments around the world are providing general guidance on the responsible creation and use of AI.[8–14] The key aspects can be categorised into three main groups:

- accountability (non-discriminatory and fair);
- responsibility (human agency and oversight); and
- transparency (technical robustness and data governance).

Cite this as: COPE Council. COPE Discussion Document: Artificial intelligence (AI) in decision making — English. https://doi.org/10.24318/9kvAgrnJ

Version 1: September 2021.

Developing and applying AI tools without careful reflection is likely to lead to problematic outcomes and cause unintended harm. In common with any other technology, understanding and testing is necessary to achieve appropriate and effective use, wide acceptance, and the greater benefits. The providers of AI are encouraged to be transparent and accountable so that the users can in turn use the tools responsibly.

One of the main challenges for accountability is related to biases. AI algorithms are trained on large data sources that can have inherent biases, or bias can be introduced in training rules selected by the developers. It is important to be aware of these biases, and to intend to correct them, or to proactively circumvent these in the design or application of the tools. Where this cannot be achieved, transparency on the limitations and biases is crucial. Algorithms themselves should undergo model performance evaluation on a regular basis to assess technical robustness, and be re-trained where needed.

Trust in AI is a critical consideration, and the development of these technologies should include a people centred approach, mindful of the ethical and cultural context of the users of the AI, as well as the individuals on the receiving end of any AI decisions.

Ethical dilemmas

The main ethical questions that have been raised by the scholarly community with regard to the use of AI for decision making in publishing systems are around the three key aspects of accountability, responsibility, and transparency.

Important questions are:

1. **Are there processes where full technical automation is acceptable or even expected? Similarly, are there processes where AI aided decisions would be accepted or expected?**
2. **Are there processes where full automation, AI based decisions, or both, would be deemed unethical?**
3. **What information should journals provide to authors (and reviewers) about the AI tools in use at their journal? How transparent do publishers need to be?**
4. **What happens if an author or reviewer disagrees with a decision or an action made by AI tools? What types of procedures should be put in place to appeal an AI action?**
5. **Who is held accountable for AI based decisions? How can publishers ensure human rights are respected and protected?**

> *If machines engage in human communities as autonomous agents, then those agents will be expected to follow the community's social and moral norms.*
>
> *A necessary step in enabling machines to do so is to identify these norms. But whose norms?*
>
> The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, https://b.link/ethics

Cite this as: COPE Council. COPE Discussion Document: Artificial intelligence (AI) in decision making — English. https://doi.org/10.24318/9kvAgrnJ

Version 1: September 2021.

Recommendations

At this stage of the development of AI, we recommend a cautious approach with respect to its adoption. Below are recommendations on the ethical use of AI in decision making in publishing, for publishers and editors, and for authors of research articles.

Note that this document does not address concerns about the ethical creation, design, and development of AI itself; many IT companies and broader frameworks are contributing such ideas in the literature, and recommended reading can be found in the resources and reference sections.

For publishers and editors

- At this point in time in the development of AI and the evolution of mindsets, we recommend that if an editorial decision is made by AI that provides a final outcome for an article, such as acceptance or rejection, the decision should directly involve an editor. The decision cannot be made by an AI tool alone.
- AI systems should provide support for individuals to make informed choices in accordance with their role. Human oversight is key to ensuring a fair process and respecting authors' rights during the evaluation of their manuscripts. Ultimately, the publisher remains accountable for editorial decisions, both made by AI as well as human editors.
- AI powered automation to increase processing speed, validation, quality assessment, and progression of the peer review process can be used, is acceptable, and even expected in many cases, provided the outcome does not result in a decision by the AI itself on acceptance or rejection of a manuscript. For example, if an AI tool detects a figure in a manuscript including a recognisable human face without the required consent form, the issue should be raised to the attention of the editor to make a decision on rejection, or the automation could proceed to automatically send a message requesting clarification, or the relevant documentation, from the authors.
- Misconduct and research integrity evaluations leading to expressions of concerns, retractions, or contacting researchers' institutions, should also not rely solely on AI decision making.
- Publishers should ensure that the humans (editors, authors, and reviewers) who use the technology have trust in the results (typically through a testing phase), providing guidelines and technical support, and ensuring details on how the AI computes the recommendations is supplied. Publishers should ensure the suppliers of AI tools are transparent on how the tools were built and trained. Any known biases in the algorithms or databases used should be shared. For example, a tool providing an author's citation count can be useful for the assessment of impact, but can be affected by self-citations. Users should be made aware of the limitation, or the tool supplier or publisher should add a parameter to present self-citation count.
- Publishers should consider whether their AI tools can increase or propagate bias against various groups. Submissions should be assessed on the merits of their content and not based on race, ethnicity, gender, age, nationality, or location of the author(s). Tools trained on historical datasets might need to be corrected for biases. Publishers should watch out for biases and share any information for these tools to be updated by developers based on this feedback.

Cite this as: COPE Council. COPE Discussion Document: Artificial intelligence (AI) in decision making — English. https://doi.org/10.24318/9kvAgrnJ

Version 1: September 2021.

- Publishers should take steps to be transparent about which of their publishing processes or workflows are automated, and where AI decisions are involved. Any AI powered automation should be clearly presented to the relevant participants of the peer review process—authors, reviewers, or editors—with clarification on how the algorithm provided the result or conclusion.

For authors

- Authors can challenge an editorial decision, if they have a valid reason, in accordance with the journal's standard procedures. Challenging an editorial decision that was made by AI or based on an AI recommendation should follow the same process as any editorial decision made by a human. Whether the decision was made by AI or a human editor, the journal and publisher are accountable for the editorial decision.
- Sound arguments should be presented by the authors to the journal editors or the publisher management team to highlight the ways in which the decision making process was flawed, if treatment was unfair, or if discrimination was noticed.
- Authors also have the right to be informed about which publishing processes or workflows were automated or where AI decisions were involved.

Future developments

There is both a need and a clear intent to ensure that development of AI takes into account ethical implications as it continues to evolve to replace some traditional systems. Results from a worldwide *Arm Ltd* survey on AI use and automation showed that individuals' trust in AI decisions is increasing, especially when they see evidence that AI can outperform a human. Transparency and accountability in design will allow individuals to trust and take responsibility for AI decisions. A major weakness that has been overlooked and is now being addressed is the availability, quality, and consolidation of training data for AI, as well as the concern of perpetrating unfair biases found in the datasets used. How quickly tools have to be re-trained on updated datasets or become obsolete is also a significant problem.

As technology continues to evolve, more AI tools will, by design, integrate ethical reasoning abilities as part of the human behaviours they attempt to mimic. We will continue to monitor developments, and the response and questioning of the research community, around this topic, to regularly update our recommendations on the ethical use of AI in the decision making activities relevant to the world of publishing.

Cite this as: COPE Council. COPE Discussion Document: Artificial intelligence (AI) in decision making — English. https://doi.org/10.24318/9kvAgrnJ

Version 1: September 2021.

RESOURCES AND FURTHER READING

European approach to artificial intelligence—
The EU's approach to artificial intelligence centres on excellence and trust, aiming to boost research and industrial capacity and ensure fundamental rights
https://b.link/digital-strategy

Everyday ethics for artificial intelligence—
Digestible guide for developers to help reflect on ethical issues in their everyday job designing AI services
https://b.link/fundamentals

AI ethics guidelines global inventory—
A list of frameworks that try to set out the principles of how systems for automated decision making can be developed and implemented ethically
https://b.link/algorithmwatch

Elaboration of a recommendation on the ethics of artificial intelligence—UNESCO artificial intelligence: towards a humanistic approach
http://b.link/ethics-ai

Ethics guidelines for trustworthy artificial intelligence, chapter 3—Subsection of the guidelines containing an assessment list to help assess whether an AI system being developed, deployed, procured or used adheres to the seven requirements of trustworthy AI
http://b.link/altai

AI today, AI Tomorrow – Awareness, acceptance and anticipation of AI: A global consumer perspective
https://b.link/ai-survey

REFERENCES

1. **COPE Forum 11 November 2019: Artificial intelligence (AI) in decision making**
https://cope.onl/forum-ai

2. **Crossref. Similarity check**
https://b.link/crossref

3. Bucci EM. **Automatic detection of image manipulations in the biomedical literature.**
Cell Death Dis 2018;9:400
https://doi.org/10.1038/s41419-018-0430-3

4. Cabanac G, Labbé C, Magazinov A. **Tortured phrases: A dubious writing style emerging in science. Evidence of critical issues affecting established journals (preprint).**
https://b.link/tortured-phrases

5. Checco A, Bracciale L, Loreti P, *et al.* **AI-assisted peer review.**
Humanit Soc Sci Commun 2021;8:25
https://doi.org/10.1057/s41599-020-00703-8

6. Golden J. **AI has a bias problem. This is how we can solve it. World Economic Forum**
https://b.link/human-bias

7. Omowole A. **Research shows AI is often biased. Here's how to make algorithms work for all of us. World Economic Forum**
https://b.link/ai-machine

8. **Institute of Electrical and Electronics Engineers. Ethics in action in autonomous and intelligent systems**
https://b.link/ethics-action

9. **IBM Institute for Business Value. Advancing AI ethics beyond compliance**
http://b.link/advancing-ai

10. **House of Lords Select Committee on Artificial Intelligence, report of session 2017–19. AI in the UK: Ready, willing and able?**
https://b.link/ai-committee

11. **The Government of Canada. Responsible use of artificial intelligence (AI)**
http://b.link/responsible-ai

12. **European Commission High-Level Expert Group on Artificial Intelligence. Ethics guidelines for trustworthy AI**
https://b.link/trustworthy-ai

13. **UNESCO. UNESCO launches artificial intelligence needs assessment survey in Africa**
https://b.link/ai-africa

14. Benjamins R. **A choices framework for the responsible use of AI.** *AI Ethics* 2021;1:49–53
https://doi.org/10.1007/s43681-020-00012-5

Cite this as: COPE Council. COPE Discussion Document: Artificial intelligence (AI) in decision making — English. https://doi.org/10.24318/9kvAgrnJ

Version 1: September 2021.

AUTHOR CONTRIBUTIONS

Conceptualisation:
The original forum discussion on Artificial intelligence (AI) in decision making in 2019 was conceptualised and hosted by Heather Tierney (**COPE** Council Member). The resulting discussion document was elaborated and written by Marie Soulière (**COPE** Council Member) on behalf of **COPE** Council.

Writing – original draft preparation:
Marie Soulière

Writing – review and editing:
Marie Soulière, Sonja Krane, Catriona Fennell, Howard Browman

ACKNOWLEDGEMENTS

Mattia Albergante, Nancy Chescheir, Sarah Elaine Eaton, Kim Eggleton, Suzanne Farley, Paul G Fisher, Alexander Laughery, Ana Marusic, Raymond Soulière, Siri Lunde Strømme, and Laura Wilson reviewed and provided suggestions for revisions to the document.

Cite this as: COPE Council. COPE Discussion Document: Artificial intelligence (AI) in decision making — English. https://doi.org/10.24318/9kvAgrnJ

Version 1: September 2021.

COPE *provides leadership in thinking on publication ethics and practical resources to educate and support members, and offers a professional voice in current debates*

C O P E

publicationethics.org

Registered charity No 1123023
Registered in England and Wales, Company No 6389120
Registered office: **COPE,** New Kings Court,
Tollgate, Chandler's Ford, Eastleigh, Hampshire,
SO53 3LG, United Kingdom

facebook.com/publicationethics

@CØPE LinkedIn

PROMOTING INTEGRITY IN SCHOLARLY RESEARCH AND ITS PUBLICATION