Solution Session

AI, plagiarism and text recycling: information resources for academic authors

Aysa Ekanger

Organization(s): UiT The Arctic University of Norway

Abstract

For the last few years generative AI has been worrying journal editors: what uses of generative AI should be allowed in scientific articles, how should these uses be declared at submission and on publication, and what are the associated copyright and ethical considerations?

Challenges brought about by AI come on top of older issues that editors are more familiar with, and that also may involve copyright and ethics: namely plagiarism and text-recycling (so-called self-plagiarism).

This solution session looks at AI use in scholarly manuscripts – alongside the issues of plagiarism and text recycling – in the light of copyright and ethics (research integrity). Due to time constraints, the session will focus on reuse of text (discussion of, for example, image reuse and generation will complicate the session). The objective of the session is to create informational resources that scientific journals can use to help their potential authors avoid plagiarism and problematic uses of AI and text recycling. The resources will be directed at academic authors and will take the form of flowcharts, checklists, or FAQs – depending on the choice of the session participants.

This approach may help authors distinguish problematic uses of generative AI more easily against the backdrop of more familiar issues of plagiarism and text recycling. It can also help remove the persisting misunderstandings that academic authors (and editors) still have about plagiarism and text recycling.

The session will start with a 15-minute presentation that will briefly introduce the concepts of plagiarism and text recycling, and provide a few examples of how generative AI may be used in the production or editing of text for scholarly publications. The presentation will give a brief overview of when these different ways of text reuse constitute an ethics violation (for plagiarism the answer is “always”), when they can also constitute copyright infringement – and whether open licenses or public domain status alleviate any of the issues.

During the second part of the session (35 minutes) participants will work in groups of 2-3 people (due to the room’s parliamentary layout). Each group can choose what informational resource they prefer to work on: a flowchart, a checklist or a

FAQ section that helps submitting authors detect and fix problematic areas in their manuscript. (The same resources can also be used by editors, when they evaluate submitted manuscripts.) The choice of resource may depend on the group's expertise level: e.g. a flowchart may require a deeper knowledge of the issues. During their work, the groups will have access to information cards (e.g. about the attribution condition in Creative Commons licenses, proper citation techniques, or accepted types of text recycling). Another resource that the groups will be provided with will be a few hypothetical cases describing problematic manuscripts. The groups are expected to document their results in the collaborative document during this session part.

The remaining 10 minutes of the session will be dedicated to the discussion of the challenges the groups encountered in the process of working on these resources, and on the advantages and disadvantages of this approach to addressing the issues of plagiarism, text recycling and use of generative AI.

The groups' outputs (as provided or in modified form) may be used by journals to inform their authors. The workshop itself can be a model for the training of editors and publishing support staff.

In order to have a productive session, it is expected that participants are familiar with the basics of copyright (such as economic and moral rights, term of protection, exceptions and limitations to copyright such as the right to quote) and open licenses (e.g. licenses from Creative Commons).