ACL 2023 Policy on AI Writing Assistance

Text generation models have been long available, and they are powering many existing tools assisting with input or the linguistic form of the text, like predictive keyboards or language checkers. However, the latest generation of models, exemplified by chatGPT and Galactica, is widely presented as something that handles both language and content: something that can produce long stretches of text of sufficient quality to serve as drafts of the user's own work. This development is prompting schools, journals and conferences (including ICML (https://icml.cc/Conferences/2023/llm-policy)) to update their authorship policies to address this trend.

Since these tools come from our own field, we are in the best position to appreciate their potential problems, including errors (https://dl.acm.org/doi/abs/10.1145/3571730) in the model output and potential plagiarism (https://openreview.net/forum?id=TatRHT_1cK) of the sources in the model's training data. At a conference, the reviewers donate their time as volunteers, and they may wish to be assured that they are not expected to extra check for such problems. Furthermore, there is the authorship issue: ACL generally expects the content of its submissions to be original, unpublished work of named authors or acknowledged contributors. Per ACM (https://www.acm.org/publications/policies/plagiarism-overview) definition of plagiarism, it includes not only verbatim or near-verbatim copying of the work of others, but also intentionally paraphrasing portions of another's work.

In consultation with the ACL exec, ACL 2023 expands the mandatory Responsible NLP Checklist (https://aclrollingreview.org/responsibleNLPresearch/) developed at NAACL 2022 by one more question concerning the use of writing assistants. If such tools were used in any way, the authors must elaborate on the scope and nature of their use. **Like the other questions on providing code, data, compensating any participants and obtaining IRB approvals, this question is not meant for automatic desk-rejections.** The purpose of this question, just like all others, is author reflection and establishment of research norms in the field. The authors' answers to all questions in the checklist will be disclosed to the reviewers, who will then be free to flag a paper for case-by-case ethics review if they see a problem. In an effort to further improve transparency of NLP research for the general public, this year the authors' answers to the responsible NLP checklist will also be made public as appendices to accepted papers, similarly to the reporting summaries (https://www.nature.com/documents/nr-reporting-summary-flat.pdf) published by Nature.

Here is our take on some cases frequently discussed in social media recently:

- **Assistance purely with the language of the paper.** When generative models are used for paraphrasing or polishing the author's original content, rather than for suggesting new content - they are similar to tools like Grammarly, spell checkers, dictionary and synonym tools, which have all been perfectly acceptable for years. If the authors are not sufficiently fluent to notice when the generated output does not match their intended ideas, using such tools without further checking could yield worse results than simpler-but-more-accurate English. The use of tools that only assist with language, like Grammarly or spell checkers, does *not* need to be disclosed.

- **Short-form input assistance.** Even though predictive keyboards or tools like smart compose in google docs are also powered by generative language models, nobody objected to them, since hardly anyone would try to use them to generate a long, unique and coherent text: it would simply not be practical. Similarly to language tools above, the use of such tools does *not* need to be disclosed in response to the writing assistance question.

- **Literature search.** Generative text models may be used as search assistants, e.g. to identify relevant literature. However, we expect the authors to read and discuss such references, just like the references identified by a regular search engine or a semantic literature recommendation tool. The usual requirements for citation accuracy and thoroughness of literature reviews apply; beware of the possible biases in suggested citations.

- **Low-novelty text.** Some authors may feel that describing widely known concepts is a waste of their time and can be automated. They should specify where such text was used, and convince the reviewers that the generation was checked to be accurate and is accompanied by relevant and appropriate citations (e.g., using block quotes for verbatim copying). If the generation copies text verbatim from existing work, the authors need to acknowledge all relevant citations: both the source of the text used and the source of the idea(s).

- **New ideas.** If the model outputs read to the authors as new research ideas, that would deserve co-authorship or acknowledgement from a human colleague, and that the authors then developed themselves (e.g. topics to discuss, framing of the problem) - we suggest acknowledging the use of the model, and checking for known sources for any such ideas to acknowledge them as well. Most likely, they came from other people's work.

- **New ideas + new text:** a contributor of both ideas and their execution seems to us like the definition of a co-author, which the models cannot be. While the norms around the use of generative AI in research are being established, we would discourage such use in ACL submissions. If you choose to go down this road, you are welcome to make the case to the reviewers that this should be allowed, and that the new content is in fact correct, coherent, original and does not have missing citations. Note that, as our colleagues at ICML point out (https://icml.cc/Conferences/2023/llm-policy), currently it is not even clear who should take the credit for the generated text: the developers of the model, the authors of the training data, or the user who generated it.

A separate, but related issue is use of generative models for writing code. ACL submissions may be accompanied by code, which counts as supplementary materials that the reviewers are not obliged to check and consider, but they may do so if they wish. The use of code assistants such as Copilot is also a relatively new practice, and the norms around that are not fully established. For now, we ask the authors to acknowledge the use of such systems and the scope thereof, e.g. in the README files accompanying the code attachments or repositories. We also ask the authors to check for potential plagiarism. Note that the Copilot in particular is currently the subject of a piracy lawsuit (https://www.techradar.com/news/microsoft-is-being-sued-over-github-copilot-piracy), and may have suggested snippets of code with licenses incompatible with yours. The use of code assistance does not obviate the requirements of authors to ensure the correctness of their methods and results.

Update: our policy on AI assistance in writing reviews can be found here (https://2023.aclweb.org/blog/review-acl23/#faq-can-i-use-ai-writing-assistants-to-write-my-review).

– ACL 2023 Program Chairs

🏷 Tags: | announcement | | policy |

📁 Categories: | blog |

📅 Updated: January 10, 2023