STM

GENERATIVE AI IN SCHOLARLY COMMUNICATIONS

DEC 2023

Ethical and Practical Guidelines for the Use of Generative AI in the Publication Process

www.stm-assoc.org

TABLE OF CONTENTS

INTRODUCTION

From the introduction of the printing press to the digital age, scholarly publishers have always been at the forefront of technological innovation. They were the first publishers to go digital, and continue to incorporate new technologies, such as machine learning and artificial intelligence (AI)—increasingly including large language models (LLMs) and generative AI tools—to enhance access, quality, and integrity and to accelerate innovation. In 2021, STM published a white paper that discusses how STM publishers contribute to the ethical and trustworthy development, deployment, and application of AI. That paper addresses both the potential opportunities and risks concerning the application of AI within the scholarly ecosystem and defines best-practice principles for an ethical, trustworthy, and human-centric AI. While the principles in that paper still hold, more specific opportunities and threats have emerged in relation to generative AI (GenAI).

The emergence of GenAI has significantly increased the need to carefully consider the use of AI in scholarly publishing - with respect to both its potential and its risks. GenAI can improve the quality of published material, break language barriers, and realize speed and efficiency improvements for all stakeholders in the publishing process. It can also blur the line between machines and researchers, fact and fiction (in the form of so-called hallucinations), and allows people to fake research results and outputs on an unprecedented scale. An additional challenge is that current integrity-checking tools are not adapted to GenAI, and it may be difficult to develop that capability.

Serious questions exist regarding intellectual property rights, confidentiality, privacy, and security. For example, concerns have been raised about whether intellectual property rights were respected in the building and training of the LLMs that underlie GenAI tools.

Litigation regarding the use of copyrighted works in training data for these tools is ongoing. Significant questions have arisen regarding how the submission of information to GenAI tools, e.g., in the form of prompts, respects laws and conventions around confidentiality and security.

In reaction to these developments, various initiatives have been launched since the public release of GenAI tools (most notably, the public release in October 2022 of ChatGPT powered by GPT3 and successive models). In August 2023, a coalition of organizations, of which STM is a part, released Global Principles of Artificial Intelligence. These principles concern the governance of the development, deployment, and regulation of AI systems and applications and cover issues related to intellectual property, transparency, accountability, quality and integrity, fairness, safety, design, and sustainable development. Individual publishers have also developed their own policies on how researchers, editors, and readers might use GenAI appropriately and ethically and are exploring additional ways that GenAI might be used in their workflows.

Developments in GenAI are highly dynamic and changing with unprecedented speed. For this reason, predicting how this technology will progress is difficult. This document reflects STM recommendations for best practices in the use of GenAI in the publication process that are based on the state of the technology at the time of its completion and, it is hoped, reflects general principles that will remain relevant in the near term. Drawing from guidelines that have been developed by individual publishers as well as current understanding of the legal, ethical, and regulatory issues around GenAI, these recommendations are intended to help individual publishers develop policies for the use of GenAI by authors, editorial teams, reviewers, vendors, third-party providers, and publishers themselves. STM expects to update these recommendations regularly because of the highly dynamic nature of these technologies.

Ethical and Practical Guidelines for the Use of Generative AI in the Publication Process

In this section, we address the potential use of GenAI by the various stakeholders involved in the publishing cycle (authors, editors/editorial teams, reviewers, and vendors), considering various aspects, including copyright, privacy, confidentiality, and ethics. The use of GenAI has potentially broad implications from all these angles. To foster a culture of responsible and ethical use of GenAI tools, all these implications must be carefully considered.

We make a distinction between publicly available GenAI platforms, such as ChatGPT, and specialized services making use of GenAI that have been specifically developed on curated, trusted content (in-house or externally provided) for the publishing industry. In the case of publicly available GenAI platforms, users should assume that the GenAI in question is one for which there is no documentation of provenance or controls over the future use of an author's inputs and responses (e.g., for (re)training the models or feedback loops).

Any use of GenAI has confidentiality, privacy, and copyright implications that should be carefully considered in advance. These are discussed in the specific examples of uses considered below.

Authors

Can GenAI be used in preparing a manuscript?

Policies differ by publisher and journal, so authors must consult the relevant guide for authors before submitting a manuscript.

In general, the following principles are recommended:

- Using publicly available GenAI as a basic tool that supports authors in refining, correcting, formatting, and editing texts and documents is permissible.
- Authors must disclose any use of GenAI that transcends those use cases so an editorial decision can be made as to its legitimacy.
- GenAI cannot be used to create, alter, or manipulate original research data and results, such as images, blots, photographs, x-rays, and measurements.
- GenAI cannot be credited as an author of a published work.

Uses of GenAI by Authors

Key Indicator	Permitted—disclosure not necessary	Disclosure necessary—permission by editorial teams	Not permitted
Basic author support tool (refine, correct, edit, and format text and documents)	✓		
Uses transcending basic author support tool		✓	
Create, alter, or manipulate original research data and results			✗
Credit GenAI as an author of a published work[1]			✗

[1] See also Tom McKay, "US Copyright Office Makes the First Move in the Battle over Generative AI and Copyright," IT Brew (Mar. 29, 2023), https://www.itbrew.com/stories/2023/03/29/us-copyright-office-makes-the-first-move-in-the-battle-over-generative-ai-and-copyright

Where using GenAI is allowed, authors must always consider copyright, privacy, and confidentiality implications before uploading text and other information to GenAI platforms for the purposes permitted. They must ensure that the necessary rights to all material to be uploaded (including third-party copyrighted content) have been obtained in advance. For example, if third-party copyrighted content is included as a block quote in an author's final text, it may infringe the third party's rights to authorize any GenAI ingestion of that content. Likewise, translating a copyrighted work creates a derivative work, and making or authorizing a translation is the exclusive right of the original copyright owner; this right must be respected in the GenAI context.

Note that the terms and conditions of public GenAI tools often permit the reuse of inputs in training, and any training data may inadvertently or intentionally show up as output from a GenAI tool without appropriate licensing messages or conditions. Authors could take actions to minimize risks, such as, for example, obtaining all rights necessary, using an AI with guarantees of confidentiality in its terms of use, or anonymizing their data before input. If such safeguards cannot be put in place, it is recommended that authors consider alternative ways to use GenAI.

In all cases, the integrity of the content generated by GenAI tools remains the author's responsibility.

Editorial Teams

Can editorial teams use GenAI to perform integrity checks, such as for copyright infringement, unauthorized reuse, paraphrasing, or plagiarism?

Publicly available GenAI platforms, such as ChatGPT, should not be used for these tasks. Uploading submitted content to these platforms can result in confidentiality, privacy, and copyright breaches and infringements because there is no way to ensure that these systems comply with publishers' norms and standards.

Some publishers are developing bespoke tools that address these issues and ensure that intellectual property and privacy are protected. These and other commercially

available tools may be appropriate for editorial use as long as such controls are in place. The owner of a journal (for example, a publisher or society) is responsible for commercial relationships with specialized GenAI services and making such services available to editorial teams. Checks will need to be done to ensure that these services have implemented appropriate measures to respect copyright laws and security, privacy, and confidentiality requirements. Terms and conditions will also need to have been agreed upon between the journal owner and the specialized service.

Can editorial teams fully rely on GenAI outcomes?

No. Like all tools, GenAI should be used only for assistance. Human oversight is always necessary.

Reviewers

Should reviewers use GenAI tools to review manuscripts?

GenAI should not be used to create a review of a paper—that should be done only by the appointed reviewer. Reviewing implies responsibilities and tasks that can be attributed to and performed only by humans.

Manuscripts under review, including supplementary material, should never be uploaded to publicly available GenAI services. That would likely breach the reviewer's agreements with the publisher, and it would create a risk of violating copyright, privacy, security, and confidentiality obligations. For the same reason, reviewers should not use publicly available GenAI services as basic author tools (e.g., to refine, correct, edit, and format text and documents). While an author may use public GenAI tools as basic author tools, we discourage their use by reviewers because confidentiality and privacy are more important at the review stage.

What should reviewers do when they suspect undisclosed use of GenAI by authors?

Reviewers should consult the publisher's policy on authors' use of GenAI. When reviewers suspect a violation of this policy, they should report it to the editor handling the manuscript as part of the review process. If GenAI appears to have been substantially involved in the work, this may be noted in the review as a factor affecting its accuracy and/or appropriateness for publication.

Vendors/Third-Party Service Providers

Can vendors use GenAI tools for their work?

Manuscripts, including supplementary material, should never be uploaded to public GenAI services, which might breach publisher policies and the terms and conditions of the GenAI service and infringe copyright, confidentiality, and other proprietary rights.

The use of publicly available GenAI for editing or formatting assistance should always be discussed with and approved by the publisher. When a publisher wishes to permit a third-party vendor or service provider to use GenAI tools for their work, the publisher should first conduct adequate review of the tool and its terms and conditions of use (such as for data privacy and security, intellectual property and confidentiality protection, and other risks) and make sure its contract with the vendor contains appropriate protection regarding these factors.

Readers

Can readers upload published manuscripts to publicly available GenAI platforms?

No, this should not be allowed, because this material might be used by the provider of these services in ways that violate the copyright inherent in that material or contravene confidentiality or privacy requirements.

To what extent should readers be informed when GenAI tools were used in the preparation of a manuscript?

When GenAI is used by authors in the writing process beyond its use as a basic author support tool, disclosure of this should be required as part of the submission (e.g., in Materials & Methods sections) in accordance with journal policy. Editorial teams will decide whether such use is legitimate and permissible on a case-by-case basis. If the decision is positive, the publisher should include a disclosure in the published article to make readers aware of the use of GenAI.

What role do publishers have in educating authors, reviewers, and readers about the use of AI?

Publishers should inform authors, reviewers, and readers about the GenAI policies that relate to their roles. For example, authors should be made aware of permitted uses of GenAI in preparing the manuscript, and reviewers should be made aware of relevant policies so they can check for potential unauthorized use. Likewise, editorial teams and readers should be informed about best-practice recommendations for using GenAI.

APPENDIX A: GLOSSARY

Glossary of terms

Artificial intelligence (AI) System is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.[2]

Generative AI (GenAI) is a type of AI system with a primary function of generating outputs across various content types that resemble human-created content.

APPENDIX B: EXAMPLES OF PUBLISHER POLICIES AND GUIDANCE

This is a non-comprehensive list of policies and guidance of selected STM member publishers. Inclusion here does not imply endorsement by STM.

AAAS: "Image and Text Integrity," https://www.science.org/content/page/science-journals-editorial-policies#image-and-text-integrity

AIP Publishing: "On the Use of AI Language Models in Scholarly Communications at AIP Publishing - AIP Publishing LLC," https://publishing.aip.org/about/news/on-the-use-of-ai-language-models-in-scholarly-communications-at-aip-publishing/

Association for Computing Machinery (ACM): "ACM Policy on Authorship," https://www.acm.org/publications/policies/frequently-asked-questions

Cambridge University Press: Summary of policy, https://www.cambridge.org/news-and-insights/news/cambridge-launches-ai-research-ethics-policy

[2]See Organisation for Economic Co-operation and Development (OECD) definition: https://oecd.ai/en/wonk/ai-system-definition-update

Elsevier: "The Use of AI and AI-assisted Technologies in Writing for Elsevier," https://beta.elsevier.com/about/policies-and-standards/the-use-of-generative-ai-and-ai-assisted-technologies-in-writing-for-elsevier

Emerald: "Emerald Publishing's Stance on AI Tools and Authorship," https://www.emeraldgrouppublishing.com/news-and-press-releases/emerald-publishings-stance-ai-tools-and-authorship

JAMA Network: "Nonhuman 'Authors' and Implications for the Integrity of Scientific Publication and Medical Knowledge, Health Informatics," https://jamanetwork.com/journals/jama/fullarticle/2801170

MDPI: "MDPI's Updated Guidelines on Artificial Intelligence and Authorship," https://www.mdpi.com/about/announcements/5687#:~:text=Authors%20are%20fully%20responsible%20for%20the%20originality%2C%20validity%2C,when%20submitting%20a%20paper%20to%20an%20MDPI%20journal

Springer Nature: "Why Nature Will Not Allow the Use of GenAI in Images and Video," https://www.nature.com/articles/d41586-023-01546-4

Taylor & Francis: "Taylor & Francis Clarifies the Responsible Use of AI Tools in Academic Content Creation," https://newsroom.taylorandfrancisgroup.com/taylor-francis-clarifies-the-responsible-use-of-ai-tools-in-academic-content-creation/

Wiley: "Generative Artificial Intelligence: Guidelines for Authors," https://onlinelibrary.wiley.com/pb-assets/assets/15405885/Generative%20AI%20Policy_September%202023-1695231878293.pdf

Conner Ganjavi et al., 2023, "Bibliometric Analysis of Publisher and Journal Instructions to Authors on Generative-AI in Academic and Scientific Publishing," https://doi.org/10.48550/arXiv.2307.11918

APPENDIX C: SELECTED READINGS AND RESOURCES

STM, "AI Ethics in Scholarly Communication," https://www.stm-assoc.org/2021_05_11_STM_AI_White_Paper_April2021.pdf

STM webinar "GPT-3 and the Future of Publishing & Academia" (Mar. 9, 2023), https://www.stm-assoc.org/events/gpt-3-and-the-future-of-publishing-academia-webinar/

STM webinar, "AI, Intellectual Property and Copyright" (June 13, 2023), https://www.stm-assoc.org/events/ai-intellectual-property-and-copyright/

"Global Principles on Artificial Intelligence (AI)" regarding intellectual property, issued September 2023 by 26 international organizations, https://www.stm-assoc.org/wp-content/uploads/FINAL-Global-AI-Principles-Formatted_9-5-23.pdf

Australian Publishers Association, "Artificial Intelligence and Publishing," https://www.publishers.asn.au/Web/Our-Work/Advocacy-Policy/Artificial-Intelligence-and-publishing.aspx?hkey=0a381ff9-4e92-4b54-bfc8-78e396e0f3f1

"IEEE CertifAIEd™," a certification program for assessing ethics of Autonomous Intelligent Systems (AIS), https://engagestandards.ieee.org/ieeecertifaied.html

Press Release, "Elsevier Takes Scopus to the Next Level with Generative AI," Elsevier, (Aug. 1, 2023), https://beta.elsevier.com/about/press-releases/elsevier-takes-scopus-to-the-next-level-with-generative-ai?trial=truel

Committee on Publication Ethics (COPE), "Authorship and AI Tools," https://publicationethics.org/cope-position-statements/ai-author#:~:text=COPE%20position%20statement&text=COPE%20joins%20organisations%2C%20such%20as,responsibility%20for%20the%20submitted%20work

UNESCO, "Ethics of Artificial Intelligence," https://www.unesco.org/en/artificial-intelligence/recommendation-ethics

AlgorithmWatch, https://algorithmwatch.org/en/

H. Holden Thorp, "ChatGPT Is Fun, but Not an Author," Science (Jan. 26, 2023), Generative AI in Scholarly Communications https://www.science.org/doi/10.1126/science.adg7879

Rebecca Lawrence & Sabina Alam, "Academic Publishers and the Challenges of AI," Social Science Space (Jan. 2023), https://www.socialsciencespace.com/2023/01/academic-publishers-and-the-challenges-of-ai/

Chris Stokel-Walker & Richard Van Norden, "What ChatGPT and GenAI Mean for Science," Nature (Feb. 6, 2023), https://www.nature.com/articles/d41586-023-00340-6

Gemma Conroy, "How ChatGPT and Other AI Tools Could Disrupt Scientific Publishing," Nature (Oct. 10, 2023), https://www.nature.com/articles/d41586-023-03144-w

"Initial Policy Considerations for Generative Artificial Intelligence | OECD Artificial Intelligence Papers," OECD iLibrary, https://www.oecd-ilibrary.org/science-and-technology/initial-policy-considerations-for-generative-artificial-intelligence_fae2d1e6-en;jsessionid=hDvLblylD8e3VES4J-F_Ba2HDLN4jFppAwmKwEbg.ip-10-240-5-100

Richard Van Noorden & Jeffrey M. Perkel, "AI and Science: What 1,600 Researchers Think," Nature (Sept. 27, 2023), https://www.nature.com/articles/d41586-023-02980-0

Deloitte, "The Current Enthusiasm for AI Adoption Is Being Fueled in Part by the Advent of Generative AI," https://www2.deloitte.com/us/en/pages/consulting/articles/generative-ai-legal-issues.html

Mohammad Hosseini & Serge P.J.M. Horbach, "Can Generative AI Add Anything to Academic Peer Review?," London School of Economics (Sept. 26, 2023), https://blogs.lse.ac.uk/impactofsocialsciences/2023/09/26/can-generative-ai-add-anything-to-academic-peer-review/

STM

ABOUT STM

At STM we support our members in their mission to advance research worldwide. Our over 140 members based in over 20 countries around the world collectively publish 66% of all journal articles and tens of thousands of monographs and reference works. As academic and professional publishers, learned societies, university presses, start-ups and established players we work together to serve society by developing standards and technology to ensure research is of high quality, trustworthy and easy to access. We promote the contribution that publishers make to innovation, openness and the sharing of knowledge and embrace change to support the growth and sustainability of the research ecosystem. As a common good, we provide data and analysis for all involved in the global activity of research. | www.stm-assoc.org

For more information about this publication, please contact:

Joris van Rossum
Program Director, STM Solutions
joris@stm-assoc.org

Claudia Russo
Regional Manager Public Affairs for Europe
claudia@stm-assoc.org

Generative AI in Scholarly Communications