Guideline on the Boundaries of

AIGC

Usage in Academic Publishing 2.0 ↘

Institute of Scientific and Technical Information of China

Elsevier | Springer Nature | Wiley

Taylor & Francis | Wolters Kluwer | Cambridge University Press

CONTENTS

1 Background

In recent years, artificial intelligence (AI) technology has been developing rapidly, especially with the release of ChatGPT, the AI chatbot in November 2022. Artificial Intelligence Generated Content (AIGC) has entered the public eye and is widely used. It is clear that AI is gaining the ability to generate fluent language, making it increasingly difficult to distinguish the mass of generated sentences from human-written text. Some scientists are already using chatbots as research assistants to sort through ideas, provide feedback on their work, write code, and review the literature. The impact of AI on research paper writing, producing, and other aspects of research is growing, but it is also having a major impact on the transparency and integrity of scientific research, which has attracted enormous attention from the academic community.

The main concern of the research community is that scientists, researchers, and students may fraudulently present AI-generated text as their own or simply use AIGC to produce unreliable research results. Large Language Models (LLMs) work by learning statistical language patterns from large online text databases. However, it should be noted that LLMs may generate false and misleading information, and fail to indicate the source of the information. Without output control, the utilization of AIGC may lead to the generation of inaccurate or biased viewpoints or unreliable research results, damaging the integrity of the academic ecosystem. At the same time, AIGC can make it even more difficult to detect academic misconduct (such as plagiarism and image forgery), using current approaches.

Therefore, it is crucial to develop guidelines that clearly define the boundaries of AIGC usage in the academic community.

Currently, various national or regional policy makers, publishers, and other relevant organizations (the Committee on Publication Ethics (COPE), the International Committee of Medical Journal Editors (ICMJE), the International Association of Scientific, Technical and Medical Publishers (STM), Taylor & Francis, Wiley, Springer Nature, Elsevier, Wolters Kluwer, etc.) have engaged in the discussions of AIGC usage in academic papers and have provided relevant regulations and guidelines. Notably, the descriptions and requirements in the AI-related guidelines and normative documents issued by different organizations at different times often exhibit inconsistencies, such as CONSORT-AI for clinical trials; SPIRIT-AI for clinical trial protocols; TRIPOD+AI for predictive models, etc. Therefore, based on an extensive review and study of existing research and exploration in the industry, we are committed to establishing a framework and guidelines that outlines the fundamental principles of best practice for AI technology in academic publishing. We aim to provide a comprehensive framework for the regulation of AIGC usage within the publishing industry, the scientific community, and science and technology regulators, with a further consensus on the appropriate application of AI technology. In the meantime, as an emerging technology, the nature and usage of AI will inevitably continue to evolve. Therefore, this guideline will be updated on an ongoing and timely basis as necessary.

2 Objectives

2.1 Prevent academic misconduct and enhance research integrity governance

Taking the prevention of misuse of AIGC as the goal, education to raise the awareness of scientific integrity should be strengthened, and academic integrity governance promoted, to make sure that research activities can be conducted in an ethical manner.

2.2 Guide relevant stakeholders to reach consensus on the use of AIGC

Specify the best practices that relevant parties should follow in the preparation, research and data gathering, writing, submission, peer review, publication, and dissemination of academic journal articles, and provide detailed and standardized guidance on the appropriate use of AIGC.

3 Principles

3.1 Transparency and accountability

Transparency and accountability are the fundamental principles underlying the use of AIGC in academic publishing. In the process from academic research to publication and dissemination, all application users (including researchers, authors, peer reviewers, and readers, etc.) should be aware of and explicitly disclose the use of AIGC in their work. Application providers and technical developers should clearly disclose the data training and content sources used by the application. Transparency should include data transparency, which involve datasets, data sources, and data processing methods; in addition, the use of intellectual property and copyright information should be disclosed as well. Accountability is a shared responsibility of all the key stakeholders, including researchers, research institutions, funders, policymakers, and publishers; the establishment and clarification of the accountability standards and related information is significant.

Quality and integrity assurance is fundamental to building trust in the application of AIGC in academic research. From the design and development of algorithms, to the inputs for training AIGC, to the inputs used in practical applications, the principles of accountability and transparency should be followed, and the use of AIGC should be indicated through identifiers or feedback mechanisms to ensure that the quality and integrity of academic research is not compromised by the utilization of AIGC technology.

3.2 Privacy and security

Privacy and security are the fundamental legal principles for the use of AIGC. AIGC should be used with respect to privacy and data protection, including the assessment of data, privacy, and security impact. Data should be appropriately anonymized to protect privacy, and measures should be taken to protect data security. Authors should also respect the specific privacy and confidentiality policies of the publishing house they are submitting to. Furthermore, users of LLMs should be aware that any submitted information might be scraped and used to train the model, and they should take necessary precautions to safeguard sensitive data.

3.3 Fairness

The utilization of AIGC should be under the principle of fairness to avoid bias. As AI has the risk of replicating and amplifying bias, potential sources of bias should be carefully assessed and reviewed in the process of training data selection, algorithm design, model generation, optimization, and application. A feedback mechanism should be established to monitor, review, and correct potential biases in a timely manner. Meanwhile, AIGC can help provide services such as copyediting and language polishing to reduce such cultural or linguistic unfairness.

3.4 Sustainable development

The multidisciplinary nature of AI systems make it highly suitable for addressing global concerns, such as the United Nations Sustainable Development Goals, carbon neutrality, and so forth. It also offers opportunities for public and private organizations to improve efficiency to achieve greater environmental sustainabili-

ty and responsibility. While AI systems carry promises to benefit all of humanity, including future generations, we must also recognize the energy-intensive nature of training and using GenAI, and the environmental impact cannot be ignored. Funding and other incentives providers of high-quality data, such as publications and databases created by publishers, help to extract actionable knowledge.

Sustainable development should be a core principle of AIGC itself. To minimize duplication and waste, the utilization of AIGC should avoid over-reliance on data that may be temporarily or permanently unavailable, while the functional modules of the tool should be based on recognized standards and guidelines to ensure that data are findable, accessible, interoperable, and reusable. In this process, special attention should be given to minimizing the environmental impact of GenAI and promoting the sustainable development of technology.

4 Behavioral framework/ practice guideline

AIGC can provide assistance (services) at various stages of research and academic publishing. In order to foster a conducive research environment, to address potential issues, and to prevent/reduce misuse of AIGC, this section provides a framework to guide authors, research institutions, academic publishers, and so forth, on compliant and responsible use of AIGC.

4.1 Research and writing

This section mainly provides guidance to researchers on the use of AIGC during the research and manuscript writing prior to submission.

4.1.1 Information collection

The data provided by AIGC are generated and extracted based on big data and large language models. However,

the accuracy and authenticity of which are not assessed or verified, and researchers need to confirm the reliability of the content.

Literature research: AIGC can be used to collect reference literature based on keywords or topics, classify and review the literature, summarize the conclusions, and provide references for researchers; moreover, it can help researchers identify new sources of information and keep track of the latest developments in the research field. It should be noted that the references provided by AIGC may be fictitious or outdated. Researchers using AIGC to support their literature review must carefully review and verify the authenticity of each suggestion and reference provided, and make human-led decisions on what to include in their research.

Concept clarification: AIGC can answer some basic conceptual questions to assist researchers in structuring their chapters. However, it should be noted that AIGC provides concept clarification based on the corpus of training data, making human oversight of any AIGC an essential step.

Research on viewpoint information: AIGC can collect information from the text on the viewpoints, emotions, and sentiment tendencies of the public or experts on certain issues. Researchers need to monitor and control the viewpoint information provided by AIGC, and clean up the information provided by AIGC if necessary, to ensure that researchers use only valid, unbiased material and prevent the dissemination of false, biased, or discriminatory information.

4.1.2 Statistical analysis

In some cases, researchers have collected data but are uncertain about the best statistical analysis to test their hypotheses. Researchers can use AIGC to select the most appropriate method of analysis or statistical analysis; however, the data used should be collected from their own experiments or obtained through other legitimate means，and the results of statistical analysis should be verified by the researchers to ensure the reliability and validity.

Data analysis and interpretation: Researchers may use AIGC to interpret data, calculate statistical indicators, perform simple data analysis, and describe statistical results. However, AIGC cannot replace the researcher's own interpretation of the data.

Suggestions and guidance on statistical methods: AIGC can provide researchers with suggestions and guidance on statistical analysis based on the question and relevant knowledge. However, these suggestions and guidance are solely based on the language model and knowledge base it has learned, which may lead to omissions and inaccuracies. Therefore, researchers need to assess the feasibility of the statistical analysis suggestions provided by AIGC, evaluate them with other reliable statistical analysis and data mining tools, or seek guidance and

assistance from the subject experts to finally determine whether to accept the suggestions provided by AIGC.

4.1.3 Charting

Charting and Image Generation: Based on the characteristics of the data and the intended purpose, AIGC can recommend the most suitable type of statistical graph for the application scenario. This helps present statistical results in a clear and effective manner, allowing researchers to convey their findings more efficiently. As a result, AIGC saves time in graph creation and enhances overall writing productivity. However, images generated from experiments such as Western blot, cell technology analysis, tissue cell staining, etc., must be obtained through authentic experimental research and must not be directly generated by AIGC. The use of generative AI or AI-assisted tools in the production of artwork such as for graphical abstracts is also not permitted. The use of generative AI in the production of cover art may in some cases be allowed, if the author obtains prior permission from the journal editor and publisher, can demonstrate that all necessary rights have been cleared for the use of the relevant material, and ensures that there is correct content attribution.

Statistical chart and diagram format processing and optimization: When processing and optimizing statistical charts, researchers can use AIGC tools to help adjust chart styles according to the amount of data and specific needs. These adjustments include, but are not limited to, adjusting font sizes, adding or modifying data annotations, adding legends, changing colors, etc. These optimizations are designed to make the charts clearer, more beautiful, and more readable and easy to understand. At the same time, researchers should ensure that when using the AIGC tool to adjust the chart, they follow the standard operation to ensure the authenticity and accuracy of the data.

Image format processing and optimization: This entry covers types of images including video and animation (such as video stills), photography, scientific diagrams, photo illustrations and other collages, as well as editorial illustrations such as drawings, cartoons, or other 2D or

3D visual representations. It is not acceptable to enhance, obscure, move, remove, or introduce a specific feature within an image. Adjustments of brightness, contrast, or color balance are acceptable if and as long as they do not obscure or eliminate any information present in the original. Similarly, amendments such as adding legends and image annotations are also acceptable. When researchers create or change image content using AIGC tools, they are required to provide a clear description of the content that was created or altered, an explanation of how the AI or AI-assisted tools were used in the creation or alteration process, and the name of the model or tool, version and extension numbers, and manufacturer.

Case Demonstration: Application of AIGC in Chart Creation

Example 1: Trend Analysis

◆ Scenario: Need to show the development trend of a technology over the past decade.

◆ AIGC Application: AIGC tools can generate a line chart based on the provided data, automatically adjusting the chart's axes, labels, and legends to clearly display the trend.

Example 2: Categorical Data Analysis

◆ Scenario: Need to compare the market shares of different products.

◆ AIGC Application: AIGC tools can generate bar charts or pie charts based on categorical data, recommending the most suitable format for data presentation, and automatically setting chart colors and labels for easier interpretation.

Example 3: Enhancing Charts in Academic Reports

◆ Scenario: Need to present research findings in an academic report, but the chart styles are not visually appealing.

◆ AIGC Application: AIGC tools can adjust the chart's font size, colors, data labels, and legend positions, making the charts more attractive and easier to understand.

Example 4: Brightness and Contrast Adjustment

◆ Scenario: Need to display photos of an excavation site, but the photos are too dark due to poor lighting conditions.

◆ AIGC Application: AIGC tools can adjust the bright-

ness and contrast of the photos, making the details clearer without losing the original information.

Example 5: Annotating Medical Images

◆ Scenario: Need to present a patient's medical images at a conference, highlighting some important features.

◆ AIGC Application: AIGC tools can add legends and annotations to the images, pointing out key anatomical structures or lesions to help the audience better understand the image content.

4.1.4 Text writing

In the writing process, AIGC can be used as a reference for researchers to improve the readability of the text, clarify the logic of the content, and recommend sentence patterns, and so forth. Also, AIGC can be used to summarize other scholarly publications during the research process or to generate the literature review section. However, it should not be used to generate research hypotheses, write the entire text, interpret data, or draw scientific conclusions. Due to the risk of inaccuracies, bias and omissions, summaries for research and literature review purposes should always be thoroughly checked against the original publication to ensure accurate reporting of the ideas, methodologies, results, and conclusions. All tasks related to scientific or intellectual contributions should be carried out by the researchers themselves, especially the writing of the critical parts of the paper. The purpose of using AIGC should be to focus on how to convey the scientific knowledge generated by the authors in the most readable way.

Case Demonstration: Application of AIGC in Writing

Example 1: Editing Document Grammar and Style

◆ Scenario: Concerned about the fluency of language expression while writing a paper.

◆ AIGC Application: Use AIGC tools to check for grammar and spelling errors and improve sentence structure, making the text clearer and more fluent. For example, break long sentences into short, powerful ones or use more precise vocabulary.

Example 2: Structuring a Paper Outline

◆ Scenario: Uncertain about the overall structure of the paper and how to organize the different sections.

◆ AIGC Application: Use AIGC tools to generate a paper outline, including sections such as Introduction, Litera-

ture Review, Methods, Results, and Discussion, and provide content suggestions for each part. Researchers can use these recommendations to organize their research content, ensuring a clear and logical flow in the paper.

4.1.5 Language and copyediting

Academic language services: Language should not be a barrier to academic communication and scientific dissemination. AIGC can serve as a high-standard language reviewer, improving the readability and writing quality of manuscripts, and thus removing language barriers in the dissemination of research. At present, AIGC-supported academic language services can assist non-native English-speakers in copyediting their manuscripts to meet the submission requirements of international journals. However, researchers also need to be aware that when a manuscript is submitted (in whole or in part) to the public AIGC or private AIGC services with low data security and confidentiality standards, it may become part of a large language model training corpus. Participation in AI training should be pre-agreed with co-authors and other stakeholders such as funders or publishers before using AI tools that could potentially use the manuscript (or part of it) to that end.

4.1.6 Citation management

When dealing with content recommended by AIGC, "citation relevance" is crucial. Researchers must ensure that the cited content is relevant to the paper, including the authenticity of the citation and the cited content.

Citation format check: AIGC can verify whether the cited literature conforms to the citation format of academic papers and identify possible errors or deficiencies.

Automatic citation generation: Authors are responsible for ensuring that any citations generated by AI tools are reliable, accurate, and relevant. While AIGC can assist researchers in identifying sources of citations and automatically generate citations that conform to the citation format of academic papers based on the literature information provided by the authors, it should only

be used as an auxiliary tool. Authors must be responsible for citations and carefully check the format and content of citations to ensure that citations meet the normative requirements of academic papers.

Automatic sorting of references: AIGC can automatically sort and check the reference list according to the specified citation format, thus helping authors to perform the relevant tasks of standardized citation in academic writing, saving the effort of manual work, improving the quality of the paper and the efficiency of researchers.

4.2 Submission

AIGC can assist with the submission process, but this requires the professional judgement of researchers who should bear the ultimate responsibility.

4.2.1 Authorship

Researchers must not permit AIGC tools to autonomously initiate, execute, or generate research outcomes without direct human supervision and guidance. Furthermore, AIGC tools cannot be held accountable for published works or research designs, emphasizing the necessity of human oversight and responsibility throughout the research process. Similarly, in most countries, AIGC does not have legal status or the ability to hold or transfer copyrights, which are the basic requirements for authorship. Therefore, according to the COPE position statement on AI tools, AIGC cannot perform the role of authors and cannot be listed as authors.

4.2.2 Standardized citation

All content originating from other sources must be carefully reviewed and properly cited. Authors must verify the authenticity and accuracy of the information provided by AIGC, and make reference notes to the underlying data sources, tools, collection, processing, etc.

4.2.3 Disclosure and statement

The use of AIGC should be fully and accurately disclosed. The following points should be clearly specified: the

user; the AI technology or system (with version number stated); the time and date of use; the prompts and questions used to generate the text; the parts of the text written or co-written by AIGC; the ideas in the paper generated by AIGC. If any part of the manuscript was written using such tools, this must be described in the Methods or Acknowledgments section in an open, transparent, and detailed manner.

[Template]

Statement: In preparing this paper, the authors used [name of specific AIGC tool/service] for [purpose of use: such as literature review/data analysis/charting, etc.]. The date and time of using this tool/service is [specific date and time] and the relevant text is generated with [Tips and Questions]. After using this tool/service, the authors have reviewed and edited the content as necessary and take full responsibility for the content of the publication.

4.2.4 Peer review

In the peer review process, authors and reviewers decide whether to use the AIGC tool in accordance with the policies of the reviewing client (e.g., journal or publisher). The AIGC tool may not be used if expressly prohibited by journal policy. If not expressly prohibited by journal policy, authors may use AIGC to assist in responding to review comments, while still taking full responsibility for the content of their responses. In case of doubt, the author should be able to justify his response.

In the peer review process, AIGC can categorize and label review comments, helping authors to quickly understand the comments, and recommending appropriate responses for authors efficiently address the review comments. It should be noted that the review report is considered a confidential document and the authors are not allowed to upload the review report to the AIGC during the confidential review process of the journal. In addition, the author shall prevent the leakage of the review content, and take necessary remedial measures

in a timely manner in case of information leakage.

Under the condition that it is permitted by journal policies, peer reviewers may use AIGC responsibly to extract critical information to inform the review, or to assist in drafting review comments; However, reviewers shall not uploaded the unpublished submissions to AI systems for any reason. AI can only be used responsibly to improve reviewers' comments, not to process the content being reviewed. Reviewers or reviewers should ensure that their review comments are accurate, truthful, and impartial, and check that their review comments reflect their professional views in a complete and specific way.

4.3 Post-publication /publishing
4.3.1 Data storage and sharing

On the premise of meeting relevant requirements, authors are encouraged to make their original data publicly available to control and prevent problems such as data contamination and falsification.

Authors should conduct a rigorous review of the data in their papers to ensure the accuracy, completeness, and reliability of the source data. The focus of the data review should include the methods of data collection and processing, the experiments, the accuracy and precision of measurements, and the methods of data storage, and so forth.

For research papers involving experimental process, researchers should record the experiments and the data collection process in a timely, accurate, and detailed manner to avoid errors or omissions, and submit the experimental data and process records together.

4.3.2 Submission and archiving of AIGC-related materials

Researchers are encouraged to share their research data (in certain situations), including but not limited to: original data, processed data, software, algorithms, protocols, methods, materials, etc. In particular, AIGC-generated content such as text, images, programs, and so forth, should be submitted and archived as supplementary material.

4.3.3 Detection and identification of AIGC-generated content

Journals and editors are advised to pay attention to security when using new tools to detect and identify content generated by AIGC to prevent information leakage and misuse of data. When setting up a review process for AIGC tools, in addition to developing appropriate technologies, manual review, and automated detection, it is important to be particularly vigilant about possible false positives. The detection results should be used as an auxiliary supporting basis in a comprehensive evaluation together with the scope of the paper, the requirements of the journal, and the overall quality of the paper, and so forth.

Researchers must provide clear disclosure and statement when using AIGC to generate manuscript text and other materials, otherwise it constitutes academic misconduct. For example, extracting newly generated text from AIGC and using it as manuscript content without proper citation will be considered misconduct.

5 Case Analysis

5.1 Case 1: Author blames ChatGPT for ethics and integrity concerns

Case Description: During the peer review process, the reviewers raised several concerns about the manuscript, specifically questioning whether there were misrepresented citations and whether non-existent sources had been cited. In the subsequent inquiry with the authors, they admitted to using ChatGPT to write the manuscript. However, they had not fact-checked the content nor verified the sources cited by ChatGPT. Given these issues, the authors requested an opportunity to revise the manuscript, offering to remove all content generated by ChatGPT and rewrite the article. However, this request was denied for the following reasons: the authors did not disclose the use of ChatGPT in their initial submission, and they failed to fact-check all the content. All listed authors are jointly responsible for the integrity of the article, and this oversight constitutes serious academic misconduct. Consequently, the manuscript was rejected for publication.

5.2 Case 2: The Retraction of a Medical Team's Paper

Case Description: On April 5, 2024, a medical team published a paper in a medical journal. The study claimed that alkaline water could reduce pain and alleviate symptoms in patients with chronic gouty arthritis. However, just three months later, the paper was retracted due to multiple serious issues[1].

Potential Errors by the Authors:

1. Use of AI-generated Content:

◆ Chart Errors: AI tools like ChatGPT were used to automatically generate images. Reviewers found errors in the number of bones depicted and meaningless labels such as "chlsinkestead atlvs no active greedis" and "Aliainine jerve sreiter."

◆ Text Generation: The introduction was found to be 100% AI-generated, severely impacting the accuracy and professionalism of the content.

2. Data and Reference Issues:

◆ Unverifiable References: Some references could not be found inPubMed or Google Scholar, questioning the

---

[1] Retracted paper: https://journals.lww.com/md-journal/fulltext/2024/07120/assessment_of_the_efficacy_of_alkaline_water_in.91.aspx.

credibility of the study.

◆ Unreasonable Statistical Data: Commenters pointed out that some statistical analyses were impossible, and the pain score data in the tables showed signs of fabrication.

3. Author Identity Verification:

◆ Non-institutional Email: The authors used an email address not affiliated with their institution, raising doubts about their identity's authenticity.

Outcome: the journal issued a retraction notice, withdrawing the research paper by the team. The journal stated that the retraction was due to serious concerns about data integrity and accuracy. Additionally, the journal announced ongoing improvements to its editorial review processes to prevent similar incidents in the future.

6 Conclusion

AIGC is an emerging concept involving multiple actors with diverse application scenarios in academic publishing, and there are many gray areas regarding the boundaries of AIGC application. Institute of Scientific and Technical Information of China (ISTIC), in collaboration with international Publishers Elsevier, Springer Nature, Wiley, Taylor & Francis, Wolters Kluwer, Cambridge University Press and after seeking advice from relevant parties, proposes suggestions for the use of AIGC technology in the form of principles and behavioral framework/practice guideline to prevent academic misconduct, strengthen integrity governance, and guide stakeholders to reach a consensus on the use of AIGC.

It is worth noting that the purpose of this guideline is to provide an exploratory framework for the application of AIGC technology, and there are still some more specific and practical issues that require further research. For example, the application of AIGC technology in academic publishing involves many stakeholders, including not only the owners and users of AIGC technology development and application, but also various responsible parties in the academic publishing process such as authors, journals, editors, reviewers, disseminators, audiences, and research regulators. It is of great importance to clarify the relationships between these parties, to identify key responsibilities and constraints, and to clearly define their responsibilities. On the other hand, there is no consensus on how to deal with the misuse of AIGC technology, and it is crucial to propose suggestions

for the proper handling of various misuses for the benefit of the implementation of responsibilities.

In the second edition, we have comprehensively revised and expanded the guideline, with a core highlight being the addition of case study analysis, which aims to deepen understanding through existing practical cases. Moreover, adjustments and optimizations have been made to crucial stages in the authors' paper writing process, such as standardizing the production of figures and tables, as well as the writing of text, and the peer review process after submission. These changes are closely aligned with the latest developments and advancements in science and technology and academia, aimed at enhancing the practicality and forward-looking nature of the guidelines, and assisting authors in better adapting to and leading the trends of academic research.

AIGC technology and tools are still under continuous innovation and development, and the scope and behavioral framework of this guideline will also need to be regularly adapted to meet new challenges and address emerging issues. We actively invite all parties to provide suggestions and feedback to update this guideline on the use of AIGC in order to provide stakeholders with a more specific and detailed practical framework for preventing academic misconduct.

We are aware that individual publishers' policies and practices may differ from each other, each tailored to

their unique policy and business environment as well as user scenarios. Therefore, we'd like to point out that this Guideline does not prevail over or replace any policies or guidance of individual publishing houses.We hope the Guideline serves as a reference tool for the research community, and we do encourage authors, editors and reviewers to always familiarize themselves with the polices as documented on the website of the individual publisher/journal in their work.

7 Acknowledgements

中 国 科 学 技 术 信 息 研 究 所

INSTITUTE OF SCIENTIFIC AND TECHNICAL INFORMATION OF CHINA

Institute of Scientific and Technical Information of China (ISTIC)
Scientometrics and S&T Evaluation Center

Fuxing Road 15, Beijing, China, 100038

Contact: Dr. Wenwen Zheng

Tel: +86 10-58882604

Email: zhengww@istic.ac.cn