UNIVERSITAS RIVISTA

VOL. 1 - No. 1
December 2025

Implications of AI in Academic Writing: Guidelines and Institutional Responses

Luca Lantero, *European Public Law Organization (EPLO)*
Giselle Heleg, *European Public Law Organization (EPLO)*
Aitzhan Kulumzhanova, *Eurasian National University and European Public Law Organization (EPLO)*

https://doi.org/10.65158/ROTG9710

| Abstract

Artificial Intelligence (AI) tools, particularly Large Language Models (LLMs) like ChatGPT, are transforming scholarly writing. Researchers and educators are struggling to determine how to integrate these tools responsibly while upholding academic integrity. This paper[1] examines the implications of AI-assisted writing, ethical issues surrounding its use, and the responses of academic institutions and journals. It presents clear guidelines for responsible AI use in scholarly writing and analyses how universities, journals, and international organizations are formulating policies. To this end, we propose a typology of AI interventions in writing and a framework for transparency to standardize disclosure.

The discussion is grounded in editorials from leading journals, policies by publishers, and guidance from organizations. Concrete recommendations are offered for stakeholders to navigate the rapidly evolving landscape of AI in academic writing.

Keywords: Artificial Intelligence (AI), Ethics, Integrity, Intellectual Property, Research, Transparency

1. *Introduction*

The rapid advancement of artificial intelligence (AI)[2], particularly in the domain of Natural Language Processing (NLP)[3] and generative language models, is reshaping the landscape of scholarly

[1] This article was developed in accordance with "AI Assessment Scale (AIAS) Level 2 – AI for Ideation and Outlining", as defined in the typology proposed herein. AI tools were used during the early stages of ideation and planning – for example, to explore potential subtopics and organize section structure. However, all paragraphs, arguments, and final text were written entirely by the author. The author retained full responsibility for content development, source interpretation, and scholarly framing.

[2] The Organisation for Economic Co-operation and Development (OECD) defines artificial intelligence within the *Explanatory Memorandum on the updated OECD definition of an AI system* as «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». OECD, *Explanatory Memorandum on the OECD Framework for the Classification of AI Systems and the Updated OECD Definition of an AI System*, Paris, OECD, 2024.

[3] Natural Language Processing (NLP) refers to the capability of a machine to process, analyse, and simulate human language – whether spoken or written – enabling interaction and interpretation in ways that resemble human communication EU–US Trade and Technology Council (TTC), *Terminology and Taxonomy for Artificial Intelligence*, Annex to the Joint Roadmap on Evaluation and Measurement Tools for Trustworthy AI and Risk Management, European Commission & U.S. Government, 2024.

communication. Tools such as OpenAI's ChatGPT, Google's Gemini, Meta's LLaMA, xAI's Grok, DeepSeek, and Anthropic's Claude have brought sophisticated writing capabilities within reach of the global academic community enabling users to generate text, summarize sources, structure arguments, and revise prose with unprecedented ease (Bubeck et al., 2023)[4].

These developments, while offering significant potential to enhance productivity and inclusivity in academic writing, also present complex ethical, epistemological, and institutional challenges that remain underexamined in scholarly discourse[5].

Historically, the evolution of scholarly communication has been marked by transformations driven by technological innovation: from the printing press to digital publishing, and more recently, to open access and preprint servers. Each of these transitions has prompted the academic community to revisit questions of authorship, authority, and access. The rise of generative AI represents a similarly disruptive shift – one that compels institutions, researchers, journal editors, and policymakers to confront urgent questions about originality, transparency, and academic integrity[6]. Unlike prior tools that merely supported the writing process, generative AI can actively shape, suggest, or produce textual content. This blurring of boundaries between tool and co-creator raises fundamental concerns about the provenance of ideas, the authenticity of scholarly voice, and the accountability of authors.

Initial responses from universities, academic publishers, and international organizations reveal a variety of positions. Some have issued permissive guidelines focused on disclosure and oversight, while others have adopted restrictive policies equating AI-generated content with plagiarism or ghost-writing[7]. Journals such as "Nature", "The Lancet", "The New England Journal of Medicine" (NEJM), and the "Journal of the American Medical Association" (JAMA) have explicitly stated that AI tools cannot be credited as authors and require full disclosure of any AI use during manuscript preparation, including how and where such tools were applied[8]. Organizations such as United Nations Educational, Scientific and Cultural Organization (UNESCO), the Council of Europe (CoE) and the European Commission (EC) have begun to develop ethical principles and governance frameworks to address the use of AI in research and education. However, these efforts remain fragmented, and a unified academic standard for AI use in writing has yet to be realized.

[4] Bubeck, S., Chandrasekaran,V., Eldan, R., et al., *Sparks of Artificial General Intelligence: Early experiments with GPT-4*, arXiv preprint arXiv:2303.12712, 2023. https://doi.org/10.48550/arXiv.2303.12712.

[5] Hosseini, M., Resnik, D. B., Holmes, K., *The ethics of disclosing the use of artificial intelligence tools in writing scholarly manuscripts*, "Research Ethics", Vol. 19, no. 4, 2023. https://doi.org/10.1177/17470161231180449.

[6] Thorp, H. H., *ChatGPT is fun, but not an author*, "Science", Vol. 379, no. 6630, 2023, p. 313. https://doi.org/10.1126/science.adg787.

[7] New York University Steinhardt, *Academic integrity and syllabus support in the age of generative AI*, NYU Steinhardt, 2023. https://steinhardt.nyu.edu/faculty-and-staff/academic-affairs/steinhardt-ai-hub/academic-integrity-and-syllabus-support-age. [last access 15 September 2025]. Some institutions have adopted policies that allow the use of AI, provided it is properly disclosed and cited. For instance, New York University states that any AI-generated content used in assignments must be appropriately attributed, and failure to disclose such use may be considered plagiarism. In contrast, the University of Hong Kong (HKU) implemented a temporary ban on the use of ChatGPT and other artificial intelligence (AI) tools in all classes, assignments, and assessments. Any violation was considered plagiarism unless the student obtained prior written consent from the course instructor. CGTN, *University of Hong Kong issues interim ban on ChatGPT*, AI-based tools, 19 febbraio 2023. https://news.cgtn.com/news/2023-02-19/University-of-Hong-Kong-issues-interim-ban-on-ChatGPT-AI-based-tools-1hxWzqgcMxy/index.html. [last accessed 15 September 2025].

[8] Lau, A., *The stance of academic journals on the use of AI*, Elion, 2023. https://elion.nz/the-stance-of-academic-journals-on-the-use-of-ai/. [last accessed 15 September 2025].

This paper addresses the growing need for principled, actionable guidance on the integration of AI into academic writing. It aims to (1) establish robust ethical and methodological guidelines for researchers using AI in scholarly work; (2) analyse the implications of generative AI for authorship, originality, bias, and responsibility; and (3) assess institutional responses through a comparative review of policies adopted by journals, universities, and international bodies.

To facilitate responsible innovation, the paper proposes a typology of AI involvement in writing, as well as a framework – the AI Use Transparency Index – to support consistent disclosure practices. Through this analysis, we aim to contribute to a critical and constructive dialogue on how to preserve the integrity of scholarly communication in an era of accelerating technological change.

2. *Conceptual and Regulatory Foundations*

Recent advances in generative artificial intelligence – particularly the development of large language models (LLMs) – have significantly expanded the scope of content creation, raising pressing questions about authorship and academic integrity. Generative AI[9], in this context, refers broadly to systems designed to produce novel content – such as text, images, or audio – based on patterns identified in extensive datasets[10]. LLMs represent a prominent subclass of generative AI: these are neural networks[11] trained on large-scale text corpora to model language and generate or transform textual outputs. The term "large" denotes the substantial number of model parameters, which, combined with vast training data, typically results in enhanced performance.

In practice, these LLMs often serve as foundation models[12] with broad capabilities: once trained, a foundation model can be adapted or "fine-tuned"[13] for many specific tasks. While these systems offer unprecedented capabilities in content generation, their use in scholarly contexts does not substitute the ethical and epistemic role of human authorship. In academic writing, authorship entails that identifiable human individuals assume intellectual responsibility for the content produced, ensure the accuracy of claims, and uphold standards of integrity, including transparency and proper attribution.

[9] Generative AI refers to artificial intelligence systems capable of producing novel content – such as text, images, music, or code – by learning patterns from large-scale datasets and using this knowledge to generate outputs that resemble human-created material. These models underpin a wide range of applications, including chatbots, text-to-image generators, and advanced language models, facilitating tasks from creative writing to software development. While generative AI offers significant potential for innovation, it also raises critical concerns related to authenticity, intellectual property, and ethical use.

[10] Toner, H., *What are generative AI, large language models, and foundation models?*, Center for Security and Emerging Technology (CSET), 2023. https://cset.georgetown.edu/article/what-are-generative-ai-large-language-models-and-foundation-models/. [last accessed 15 September 2025].

[11] A neural network is a system made up of layers of connected units called neurons. It takes in data, processes it through these layers to find patterns, and learns to make predictions by adjusting connections based on mistakes it makes.

[12] Foundation models are large-scale models trained on broad and diverse datasets and designed to be adaptable across a wide range of downstream tasks. The term "foundation model" is often used interchangeably with "general-purpose AI" (GPAI). While policy and regulatory frameworks – such as the EU AI Act – tend to prefer the term GPAI, the expression "foundation model" is more prevalent in technical and research communities. For the purposes of this study, the term foundation model will be used. These models can be trained on various types of data, including text, images, speech, and 3D signals. Their capabilities include object recognition, sentiment analysis, question answering, instruction following, and image captioning.

[13] Fine-tuning is a process in machine learning where a pre-trained AI model is further trained on a specific dataset to adapt it for a particular task or domain. Fine-tuning allows organisations to customise powerful AI models for their specific needs, improving performance while saving time and resources compared to training from scratch.

Leading publication ethics bodies have clearly stated that AI tools cannot fulfil the requirements for authorship, as they lack accountability and legal standing. This position, including disclosure obligations, is examined in greater depth in the section on authorship below.

These conceptual boundaries are echoed in emerging global ethical frameworks. The UNESCO *Recommendation on the Ethics of Artificial Intelligence*[14] – the first global AI ethics instrument formally adopted by 193 Member States – places human rights, dignity, and agency at the core of AI governance. It articulates key normative principles such as transparency, fairness, accountability, and the indispensability of human oversight in the deployment of AI systems. In a complementary effort, UNESCO's *2023 Guidance for Generative AI in Education and Research* promotes a human-centered approach to technological integration in academic settings. The guidance calls on governments and institutions to «implement immediate actions, plan long-term policies and develop human capacity» to ensure that generative AI strengthens, rather than compromises, educational and research ecosystems[15].

Alongside UNESCO, the OECD's AI Principles articulate a complementary vision: they promote AI that is «innovative and trustworthy» while fully respecting human rights and democratic values[16]. These principles – now endorsed by numerous countries – emphasize the necessity of transparency and explainability, requiring AI actors to provide clear and accessible information regarding a system's capabilities, limitations, and underlying decision-making logic. This enables both users and those affected by AI outputs to understand, assess, and, if necessary, contest automated outcomes.

Accountability is likewise central to the OECD framework, which advises that developers and deployers maintain robust mechanisms for traceability. This includes documenting training datasets, design decisions, and decision pathways to facilitate external auditing and verification of AI-generated outputs. These requirements are closely aligned with UNESCO's recommendations, which call for both technical and institutional safeguards to ensure the auditability and traceability of AI systems, along with mechanisms that embed explainability and independent oversight[17].

Taken together, the guidance offered by these leading international organizations converges on a shared normative foundation – centred on transparency, accountability, traceability, explainability, and human oversight – which is essential to the ethical governance of AI, particularly in domains where the societal and ethical implications are substantial.

In the European Union, the *Artificial Intelligence Act*[18], which entered into force in August 2024, represents the first comprehensive and binding legal framework for AI at the global level. Its overarching

[14] UNESCO, *Recommendation on the Ethics of Artificial Intelligence*, Paris, UNESCO, 2021. https://unesdoc.unesco.org/ark:/48223/pf0000381137. [last accessed 15 September 2025].
[15] UNESCO, *Guidance for Generative AI in Education and Research*, Paris, UNESCO, 2023. https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research. [last accessed 15 September 2025].
[16] OECD, *Recommendation of the Council on Artificial Intelligence (OECD AI Principles)*, OECD Legal No. 0449, adopted 22 May 2019, updated 2024, Paris, OECD, 2024. https://oecd.ai/en/ai-principles. [last accessed 15 September 2025].
[17] UNESCO, *Recommendation on the Ethics of Artificial Intelligence*... cit.
[18] EU Artificial Intelligence Act, *The EU Artificial Intelligence Act Up-to-date developments and analyses of the EU AI Act*. https://artificialintelligenceact.eu. [last accessed 15 September 2025].

objective is to promote the development and deployment of "trustworthy AI" across Europe[19]. The Act adopts a risk-based regulatory approach, prohibiting certain applications deemed unacceptable – such as AI systems that manipulate human behaviour or engage in indiscriminate biometric categorization – while imposing stringent requirements on high-risk systems, including those used in education. These obligations include safeguards to protect fundamental rights, such as transparency, human oversight, and the assurance of data quality. In addition to the binding legislation, the EU has introduced voluntary instruments, such as the *AI Pact*,[20] which encourages AI developers and providers to commit proactively to these standards ahead of formal compliance deadlines.

In parallel, the *Guidelines on the Responsible Use of Generative AI in Research*[21] issued by the European Commission[22] provide structured advice for researchers and institutions on how to uphold established norms of research integrity in the context of AI-assisted work. The guidelines recommend, among other points, that generative AI tools should not be employed in sensitive processes such as peer review, that any use of such tools be thoroughly documented and transparently disclosed, and that institutions establish mechanisms to oversee and support responsible AI use within their research environments. At the level of scholarly publishing, key organizations such as the "Committee on Publication Ethics" (COPE) and the "International Committee of Medical Journal Editors" (ICMJE) have issued position statements emphasizing that generative AI tools cannot meet the conditions for authorship, as they lack legal personhood and cannot assume responsibility. These bodies also require full disclosure of any AI-generated content in submitted manuscripts[23].

In parallel, the *Council of Europe's Framework Convention on Artificial Intelligence*, opened for signature in 2024, obliges signatory states to ensure that the development and deployment of AI systems are fully aligned with human rights, democracy, and the rule of law[24]. The Council's education initiatives reinforce this commitment by stressing that AI in learning contexts must promote accessible and inclusive education, with particular attention to data privacy, algorithmic transparency, and bias mitigation.

Together, these European instruments contribute to operationalizing abstract principles into concrete expectations for academic AI use (full disclosure, human-in-the-loop review, and protection of privacy and IP in data, among others).

Outside the European context, national and regional approaches to the governance of generative AI are evolving rapidly. UNESCO has played a leading role in encouraging global cooperation to implement

[19] European Commission (EC), *Artificial Intelligence Act. Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence*, Brussels, European Commission, 2024.
[20] The AI Pact encourages and supports organisations to plan ahead for the implementation of AI Act measures. European Commission, *AI Pact*. https://digital-strategy.ec.europa.eu/en/policies/ai-pact. [last accessed 15 September 2025].
[21] The EC, together with the European Research Area (ERA) countries and stakeholders, has put forward a set of guidelines to support the European research community in their responsible use of generative artificial intelligence (AI). *European Commission, Living guidelines on the responsible use of generative AI in research*. 15 April 2025. https://research-and-innovation.ec.europa.eu/document/2b6cf7e5-36ac-41cb-aab5-0d32050143dc_en. [last accessed 15 September 2025].
[22] European Commission (EC), *Artificial Intelligence Act. Regulation*... cit.
[23] European Journal of Therapeutics. (n.d.), *Authorship and artificial intelligence (AI) tools*, Gaziantep University Faculty of Medicine, 2025. https://eurjther.com/index.php/home/Authorship-AI. [last accessed 15 September 2025].
[24] Council of Europe (CoE), *Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*, "Council of Europe Tereaty Series", Vol. no. 225, 5 September 2024. https://rm.coe.int/1680afae3c. [last accessed 15 September 2025].

ethical frameworks, calling on governments to involve diverse stakeholders and to ensure that AI development is «guided by sound scientific research as well as ethical analysis»[25].

In Asia, China introduced the *Interim Measures for the Administration of Generative AI Services* in 2023, representing one of the first regulatory frameworks explicitly targeting AI content providers. These Measures define generative AI providers as responsible entities, obligated to remove illegal or harmful content, report incidents to relevant authorities, and continuously improve their models to prevent recurrences[26]. Furthermore, providers must ensure lawful data-curation practices – respecting intellectual property and data privacy – and guarantee that training datasets are high quality and appropriately labelled. All AI-generated content must be clearly identified, for instance through tagging or watermarking, and user protections are emphasized through restrictions on unnecessary data collection and requirements for user consent regarding personal information.

Importantly, the regulatory framework in China seeks to balance innovation and control. The Measures explicitly advocate for a model of governance that encourages innovation while ensuring compliance with legal and ethical standards. Other countries in the region are similarly developing policies or draft regulations addressing AI use in education and research. Singapore, for instance, has taken steps to address the use of AI in education through the development of the *AI-in-Education Ethics Framework* (AIEd), which builds on the national Model AI Governance Framework and reflects the core values of the teaching profession[27]. Though approaches vary, a common set of principles is emerging transparency through clear labelling and disclosure, accountability through human responsibility, and robust oversight mechanisms to ensure that AI remains a tool under human control.

Taken together, these conceptual and regulatory developments shape the contemporary discourse on generative AI in academic writing, while reaffirming that intellectual control and ethical responsibility must rest with human authors.

Core values of academic integrity – particularly accountability and integrity – are operationalized through the requirement to disclose AI use and the explicit exclusion of AI systems from authorship attribution. The prevailing consensus increasingly treats generative AI as an advanced scholarly tool, whose use must be rigorously documented, critically assessed, and fully supervised by human researchers.

3. *Guidelines for Responsible AI Use in Academic Writing*

Considering the opportunities and risks posed by AI, it is critical to establish robust guidelines for its responsible use in research writing. The following principles provide a foundation to harness generative AI tools ethically and transparently:

(i) Maintain Originality and Avoid Plagiarism: All submitted academic work must remain the

[25] UNESCO, *Recommendation on the Ethics of Artificial Intelligence*... cit.
[26] Cyberspace Administration of China, *Interim Measures for the Administration of Generative Artificial Intelligence Services*, 2023.
[27] Ministry of Education, Singapore, *AI in Education (AIEd) Ethics Framework*, "Student Learning Space", Singapore, 2025. https://www.learning.moe.edu.sg/ai-in-sls/responsible-ai/ai-in-education-ethics-framework/. [last access 15 September 2025].

original intellectual product of the human authors. Using AI to generate substantial text without acknowledgment is essentially plagiarism. In fact, «a submitted academic manuscript must be the original work of the authors, and the wholesale use of ChatGPT without formal acknowledgement is akin to plagiarism»[28]. To uphold integrity, authors should only use AI to assist with ideas or wording that they fully understand and can integrate into their own original argument. Any AI-generated passages should be treated as third-party content – quoted or paraphrased and cited if reproduced verbatim – to clearly distinguish the author's contributions.

(ii) Ensure Transparency Through Disclosure: Full transparency regarding the use of generative AI tools is increasingly recognized as both an ethical obligation and a methodological requirement in scholarly writing. Authors are expected to provide explicit disclosure of whether, how, and to what extent AI technologies were employed during any stage of the research and writing process – including data analysis, drafting, or linguistic refinement. A growing consensus among leading publishers and research ethics bodies affirms that transparency is essential to maintaining the integrity of academic communication. For example, Elsevier now mandates that manuscripts include a dedicated section titled *Declaration of Generative AI and AI-assisted Technologies*, wherein authors must describe any use of AI tools, using a standard template for consistent wording[29]. Similarly, the Committee on Publication Ethics (COPE) and numerous editorial guidelines advise researchers to disclose AI usage even when its contribution is minimal, reinforcing the principle that full accountability rests with the human author[30]. Disclosures should specify the AI tool employed (e.g., GPT-4), the purpose of its use (such as linguistic refinement or drafting a methods section), and the extent of its contribution. Such transparency enables editors, reviewers, and readers to assess the provenance and credibility of the work, thereby reinforcing trust in the publication process.

(iii) Uphold Accountability and Human Oversight: The ultimate responsibility for the content of a paper lies with its human authors, and this responsibility cannot be outsourced to an AI. AI tools must not be listed as authors on papers. This guideline has been endorsed by major publishers and COPE: «AI tools cannot meet the requirements for authorship as they cannot take responsibility for the submitted work»[31]. Human authorship entails duties – conception of ideas, critical interpretation, accountability for the accuracy of data and claims, and the capacity to respond to criticism – which only humans can fulfil. Consequently, authors who employ AI are expected to exercise careful oversight: they should verify all AI-generated content for accuracy, correct any errors, and ensure that no biases or unethical content have been introduced. In practice, this means thoroughly fact-checking AI contributions against source literature and subjecting AI-assisted text to the same level of critical revision as any other source. Publishers like "Springer Nature" underscore that «what's fundamental is that there is clarity [...] We need transparency, as that lies at the very heart of how science should be done and communicated»,

[28] Nature, *Tools such as ChatGPT threaten transparent science; here are our ground rules for their use*, 2023. https://www.nature.com/articles/d41586-023-00191-1. [last accessed 15 September 2025].
[29] Elsevier, *Generative AI policies for journals*, n.d. https://www.elsevier.com/about/policies-and-standards/generative-ai-policies-for-journals. [last accessed 15 September 2025]. Elsevier, *The use of generative AI and AI-assisted technologies in writing for Elsevier*, n.d. https://www.elsevier.com/about/policies-and-standards/the-use-of-generative-ai-and-ai-assisted-technologies-in-writing-for-elsevier. [last accessed 15 September 2025].
[30] COPE, *Artificial intelligence and authorship*, 2023. https://publicationethics.org/news-opinion/artificial-intelligence-and-authorship. [last accessed 15 September 2025]. COPE, *Authorship and AI tools*, 2023. https://publicationethics.org/guidance/cope-position/authorship-and-ai-tools. [last accessed 15 September 2025].
[31] Ibid.

and that authors remain accountable for any content produced with AI assistance[32]. In sum, AI may assist with writing, but it cannot replace the author's intellectual responsibility.

(iv) Define Appropriate Scope of AI Assistance: Researchers should use AI as a support tool, not a replacement for human scholarly effort. Appropriate uses include tasks like language polishing, grammar and spelling correction, formatting assistance, or generating non-substantive text that the author then rigorously reviews[33]. Many journals permit using AI to improve readability and language of a manuscript, especially to help non-native English speakers express their ideas more clearly. For example, Elsevier's guidelines allow AI use «to improve the readability and language of the research article, but not to replace key tasks that should be done by the authors, such as interpreting data or drawing scientific conclusions». By contrast, it is considered irresponsible to rely on AI for core scholarly tasks such as analysing results, formulating arguments, or drawing novel conclusions, since these require human expertise and insight. Authors should also refrain from using AI to generate citations or factual content that they have not verified, as AI-generated references are often fabricated and factual errors can be introduced. The guiding rule is that AI may facilitate expression of the author's ideas, but it must not produce the ideas or analyses themselves.

(v) Mitigate Bias and Validate Accuracy: Any content produced with the aid of AI must be carefully checked for biases, errors, and omissions. LLMs are known to generate text that appears coherent and plausibly accurate, yet may include erroneous or fabricated information, raising concerns about their reliability in knowledge-intensive tasks. They can also inadvertently amplify societal biases present in their training data. Responsible use of AI in academic writing therefore entails rigorous critical evaluation of AI outputs. If an AI tool is used to generate a summary or paraphrase, the author should cross-check the summary against the original sources to ensure fidelity and completeness. If the tool suggests an analysis or interpretation, the author must confirm that it is logically and scientifically sound. Human judgment must remain in charge. Some have suggested a requirement for a "human in the loop" review, meaning no AI-generated text enters the final manuscript unless a human author has scrutinized and approved it. This aligns with emerging international ethics guidelines: UNESCO's *Recommendation on the Ethics of AI* explicitly states that AI systems should not displace human determination, and that humans must retain «ultimate responsibility and accountability» for decisions or content produced by AI[34].

(vi) Protect Data Privacy and Intellectual Property: Responsible AI usage also involves safeguarding confidential data and respecting copyrights. If researchers input portions of an unpublished manuscript or dataset into an online AI service, they risk violating privacy or data protection rules, as well as exposing novel data or ideas prematurely. Many university guidelines (e.g., *Harvard University's Initial Guidelines for the use of Generative AI Tools*)[35] advise never to paste sensitive, proprietary, or unpublished information into generative AI platforms without proper assurances of privacy. Likewise, using AI to generate text does not exempt authors from copyright considerations – authors should

[32] Vincent, J., *ChatGPT can't be credited as an author, says world's largest academic publisher*, The Verge, 2023. https://www.theverge.com/2023/1/26/23570967/chatgpt-author-scientific-papers-springer-nature-ban. [last accessed 15 September 2025].
[33] Singapore Management University, *Disclosure on the use of AI in research manuscripts: How are researchers doing it?*, 2024. https://library.smu.edu.sg/topics-insights/disclosure-use-ai-research-manuscripts-how-are-researchers-doing-it. [last accessed 15 September 2025].
[34] UNESCO, *Recommendation on the Ethics of Artificial Intelligence*...cit.
[35] Harvard University, *Research with Generative AI*, n.d. https://www.harvard.edu/ai/research-resources/. [last accessed 15 September 2025].

not have an AI mimic someone else's writing without attribution, and they must be aware that AI-generated text may not be copyrightable. As a precaution, several publishers (e.g. "Springer Natureù") have even forbidden AI-generated images or figures in publications due to unresolved copyright and integrity issues. Researchers should use AI tools that are compliant with data protection standards or use self-hosted models for sensitive work. By adhering to these guidelines – originality, transparency, accountability, appropriate use, bias mitigation, and privacy – scholars can leverage AI's benefits (efficiency, enhanced expression, expanded ideation) while upholding the rigor and trustworthiness of academic writing.

4. *Ethical and Institutional Implications*

As it was previously mentioned, the rise of AI in academic writing carries far-reaching implications. It challenges traditional notions of authorship, raises novel ethical questions, and necessitates new norms in scholarly communication. Key areas of concern include transparency of AI involvement, plagiarism and originality, biases and reliability of AI outputs, and changes in the research workflow.

This section explores these challenges and analyses how institutions have responded, spanning approaches from integration and experimentation to formal regulation.

4.1 *Transparency and Disclosure as New Norms*

Transparency regarding the use of artificial intelligence has become a central ethical principle in the academic response to LLMs. A primary concern is that undisclosed AI assistance may compromise the credibility of scholarly work. When sections of a manuscript are generated by AI without proper disclosure, readers and peer reviewers may be misled about the origin of ideas, the originality of the language, or the extent of the author's intellectual contribution. Moreover, insufficient transparency obstructs key pillars of academic integrity, such as reproducibility and accountability, since concealed AI-generated content cannot be properly evaluated, scrutinized, or replicated by others.

Academic journals have thus rapidly instituted disclosure requirements. In early 2023, the journal "Science" announced an updated editorial policy banning the use of ChatGPT-generated text "entirely" in submissions and requiring that all content be the original work of the authors[36]. Its editor-in-chief, at that time in fact argued that even using ChatGPT to prepare a paper risked injecting errors and shortcuts that "could be lost" in terms of context and rigor while he insisted that the proper direction for science was deeper scrutiny and not reliance on AI to summarize findings. Most other top journals did not go so far as an outright ban but converged on mandatory disclosure. "Springer Nature" clarified that it permits AI-assisted writing «as long as this contribution is properly disclosed by the authors». In January 2023, "Nature" stated its "ground rules" for AI: «no LLM tool will be accepted as a credited author on a research paper», and any use of such tools for writing or image generation must be documented[37]. This stance was echoed by other major publishers. For example, Elsevier's policy required that authors

[36] The Guardian, *Science journals ban listing of ChatGPT as co-author on papers*, "The Guardian", 2023. https://www.theguardian.com/science/2023/jan/26/science-journals-ban-listing-of-chatgpt-as-co-author-on-papers. [last accessed 15 September 2025].
[37] Nature, *Tools such as ChatGPT threaten transparent science; here are our ground rules for their use*, 2023. https://www.nature.com/articles/d41586-023-00191-1. [last accessed 15 September 2025].

«acknowledge all sources and contributors», including AI, and to document its use appropriately. "Taylor & Francis" itself issued guidance that authors using AI should do so only in accordance with publishing ethics and with explicit acknowledgment in the work.

By mid-2023, disclosing AI assistance had become a *de facto* expectation in scholarly publishing. Some universities and funding agencies similarly encourage or require disclosure in theses and reports, reinforcing that transparency is integral to research integrity. The *European Code of Conduct for Research Integrity*[38] now advises researchers to report any use of AI tools during research and writing as part of maintaining honesty and accountability. This cultural shift towards openness is intended to uphold trust: readers should never be left guessing whether a human or an AI wrote a given sentence or analysis.

The emphasis on transparency is also a response to practical realities. There is currently no foolproof method to detect AI-written text. While software detectors exist, they yield false positives and negatives, and their effectiveness can be undermined by relatively minor modifications in phrasing. Recognizing these constraints, the "International Conference on Machine Learning" acknowledged in its ban on AI-generated submissions that the rise of tools like ChatGPT presents «unanticipated consequences [and] unanswered questions», including unresolved issues around authorship, ownership, and the originality of AI-generated content. The policy prohibits text "produced entirely" by AI, while permitting its use for editing or polishing author-written material[39].

In this context, voluntary self-disclosure emerges as the most reliable and constructive mechanism for identifying AI involvement in scholarly work. Promoting a culture of transparency – rather than one of surveillance – is increasingly viewed as both more effective and more aligned with academic norms. As noted in recent European Commission guidelines, institutions should foster «an atmosphere of trust where researchers are encouraged to transparently disclose the use of generative AI without concerns for adverse effects»[40]. This approach entails both removing the stigma associated with AI-assisted writing and clearly delineating ethical boundaries to ensure responsible and accountable use.

That said, implementing transparency is not without challenges. How should disclosure be formatted? How much detail is enough? Policies vary. Some journals ask for a statement in the Methods or Acknowledgments. Others, like certain Elsevier journals, require a formal declaration section at the end of the manuscript[41]. There is also debate on whether prompts given to AI should be included for context. Since an AI's output depends heavily on the prompt, some argue that providing the exact prompt could be useful for transparency and for other researchers to understand the interaction. Additionally, consistent terminology is still emerging. Terms like "AI-assisted writing" or "generated with ChatGPT" need clear definitions to differentiate between minimal editing versus substantive content generation. Despite these complexities, the trajectory is clear: transparent reporting of AI contributions is becoming a standard part of academic integrity, much like disclosing funding sources or potential conflicts of interest.

[38] ALLEA, *The European Code of Conduct for Research Integrity*, 2023. https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/european-code-of-conduct-for-research-integrity_horizon_en.pdf. [last accessed 15 September 2025].
[39] Vincent, J., *ChatGPT can't be credited as an author...* cit.
[40] European Commission, Directorate-General for Research and Innovation, *Responsible Use of Generative AI in Research. Living Guidelines on the Responsible Use of Generative AI in Research*, ERA Forum Stakeholders' Document, Second Version, April 2025, Brussels, European Commission, 2025.
[41] Singapore Management University, *Disclosure on the use of AI in research manuscripts...* cit.

4.2 *Plagiarism, Originality, and Academic Integrity*

A key ethical concern is that AI might enable plagiarism or otherwise erode the originality of academic work. If an AI model generates text, to what extent can an author claim it as their own writing? This question strikes at the heart of academic ethics. Traditionally, plagiarism has been understood as the unacknowledged use of another person's words or ideas. In the case of AI, while the tool is not an author in the human sense, incorporating its generated content without appropriate attribution constitutes a breach of academic norms.

Leading academic journals have taken the position that undisclosed AI-generated writing may amount to plagiarism, as it involves presenting content that the researcher neither authored nor acknowledged[42] COPE reinforces this view, emphasizing that «all parts of the manuscript must be the author's original work», and that any AI-assisted contributions must therefore be transparently disclosed to uphold authorship integrity[43]. This framing situates the use of AI not as inherently unethical, but as requiring clear attribution to preserve academic honesty and accountability.

Another dimension is self-plagiarism and recycling. An author might be tempted to have AI rewrite their own prior publications in new words – this could evade plagiarism detection but would violate norms against duplicate publication. Or an AI could piece together text from various sources in its training data (some of which might be copyrighted). Early experiments showed ChatGPT can produce paragraphs that closely mimic specific articles or internet content. If authors simply copy-paste such output, they risk unintentional plagiarism of those sources. There have already been cases where AI-generated abstracts have successfully fooled peer reviewers, raising serious concerns about the potential for misuse, such as generating fake or plagiarized academic content. Considering these risks, the academic community has responded with caution. This includes banning the listing of tools like ChatGPT as co-authors and promoting clear disclosure of AI use. These actions aim to protect the integrity of scholarly publishing by encouraging transparency and discouraging unethical practices.

From the perspective of universities, AI-assisted cheating by students is a growing concern. Many academic integrity offices have revised their policies to make it clear that submitting AI-generated work as one's own, without proper citation or permission, is considered academic misconduct. This is viewed similarly to contract cheating or copying from other students. Some universities, fearing that students would use tools like ChatGPT to generate essays or exam answers, initially banned these tools from campus networks – New York City public schools, for example, implemented such a ban in early 2023[44].

In short, the community consensus is that AI must not be a shortcut to avoid the hard work of learning, researching, and writing. If misused, AI could encourage a form of academic dishonesty by enabling people to generate passable text that they never deeply engaged with. This threatens the core of scholarship, which is about developing and demonstrating understanding. The responsible path is to treat AI as just another tool – like translation software or statistical packages – that must be used in

[42] Nature, *Tools such as ChatGPT threaten transparent science*... cit.
[43] COPE, *Artificial intelligence and authorship*... cit.
[44] Meckler, L., *New York City schools ban AI chatbot ChatGPT for fear of cheating and misinformation*, "The Washington Post", 2023. https://www.washingtonpost.com/education/2023/01/05/nyc-schools-ban-chatgpt/. [last accessed 15 September 2025].

accordance with ethical guidelines. Just as copying someone's data analysis code without credit is unethical, so is copying AI prose without credit.

Conversely, if used openly and with critical oversight, AI might "reduce" plagiarism by helping writers express ideas in their own voice rather than be tempted to copy-paste from sources. It can also educate users on different ways to phrase content, potentially improving paraphrasing skills (though caution is needed to avoid inappropriate close paraphrasing). Ultimately, maintaining academic integrity in the age of AI comes down to ensuring honesty (about AI use), fairness (not gaining unfair advantage or misrepresenting one's effort), and respect (for the intellectual contributions of others and for the norms of one's discipline).

4.3 *Authorship and Intellectual Accountability*

Authorship carries not only credit but responsibility. A key ethical issue with AI-assisted writing is determining authorship and attributing intellectual responsibility. As mentioned, all major publishers have decreed that AI cannot be listed as an author on papers[45]. The rationale is straightforward: authors must fulfil certain criteria – including contributing to the work, drafting or revising it critically, approving the final version, and agreeing to be accountable for it.

AI systems do not meet the fundamental criteria required for authorship. As noted by experts, software cannot be meaningfully held accountable for a publication, cannot claim intellectual property, and is unable to engage in scholarly dialogue or clarify its contributions[46]. This position is reinforced by the COPE, which emphasizes that AI lacks legal status and agency, and therefore cannot hold copyright, be subject to liability, or formally approve a research manuscript[47]. Authorship remains a human responsibility, inseparable from the accountability it entails.

The integration of AI in the writing process has also brought renewed attention to the concept of human authorship. When researchers rely heavily on generative tools, questions arise regarding the extent to which they remain the true authors of the text. In response, academic journals have increasingly emphasized that human authors must take full responsibility for any AI-assisted content, as if they had written it themselves. "Nature", for example, considers that the use of AI is permissible only when authors assume full accountability for its output (Nature, n.d.). Similarly, "eLife" has noted that the central issue is not whether AI is used, but how its use is disclosed and managed[48]. Authors are expected to describe how AI tools contributed to the work and, by doing so, accept responsibility for the content produced. In practical terms, this means that if an AI generates a flawed or misleading passage, it is the responsibility of the human author to identify and revise it. Failing to do so places accountability on the author, who remains answerable to peer reviewers and readers for any errors, omissions, or instances of plagiarism that may result.

[45] Taylor & Francis, *Taylor & Francis Clarifies the Responsible Use of AI Tools in Academic Content Creation*, Taylor & Francis Newsroom, 17 February 2023. https://newsroom.taylorandfrancisgroup.com/taylor-francis-clarifies-the-responsible-use-of-ai-tools-in-academic-content-creation/. [last accessed 15 September 2025].
[46] Stokel-Walker, C., *ChatGPT listed as author on research papers: Many scientists disapprove*, "Nature", 2023 https://www.nature.com/articles/d41586-023-00107-z. [last accessed 15 September 2025].
[47] COPE, *Artificial intelligence and authorship*...cit.
[48] ELife *Publishing and Peer Review at eLife*, n.d. https://elifesciences.org/about/peer-review [last access 15 September 2025].

The question of "authorship attribution" extends beyond naming and includes the ordering of authors and the allocation of credit. When AI tools are used collaboratively by multiple human authors, the established human authorship order remains unaffected[49]. However, if the AI's contribution is substantial, it may be appropriate to acknowledge its use in a dedicated section (e.g., «Acknowledgments: An AI writing assistant was used to help draft portions of the introduction»). In the future, AI tools might be credited similarly to widely used software or editorial assistance. Indeed, in certain fields – where permitted – authors have already begun to mention tools like ChatGPT in the acknowledgments, treating them as non-human contributors. This practice echoes earlier acknowledgments of statistical programs or proofreading services.

What distinguishes AI, however, is its ability to generate ideas and text that may appear original or even creative. This raises concerns about intellectual provenance: omitting mention of AI tools might obscure the true origin of certain content, while attributing authorship to them risks overstating their capacity for responsibility, intention, or innovation. As a result, a balanced approach is emerging AI should be treated as a tool rather than as an author[50].

The question of intellectual accountability extends beyond the act of writing and is intrinsically linked to the responsibilities that authors bear during peer review, public engagement, and post-publication dialogue. Once a scholarly article is published, it is the human authors – not the AI tools – who are expected to respond to inquiries, clarify ambiguities, and, if necessary, issue corrections or retractions. As Magdalena Skipper, editor-in-chief of "Nature", aptly emphasised, authorship entails «responsibilities that extend beyond publication»[51], including the duty to engage with the scholarly community and uphold the integrity of the scientific record.

There is a growing concern that excessive reliance on generative AI may erode this foundation of responsibility. Authors who incorporate AI-generated content without thorough comprehension risk producing work they cannot adequately defend, thereby undermining the credibility of the research and their own scholarly authority. A useful ethical benchmark in this context is the following: can the author confidently explain, justify, and take responsibility for every statement contained within the manuscript? If sections of the work fall outside the author's expertise or awareness due to the uncritical adoption of AI-generated text, then the integrity of authorship has been compromised. Responsible academic practice requires that AI tools be employed only in ways that authors can fully oversee, validate, and articulate in response to peer or public scrutiny.

In sum, while AI may serve as a valuable support in the writing process, it cannot substitute the author's intellectual accountability. Ultimately, it is the human author's reputation, credibility, and scholarly record that remain on the line.

[49] International Committee of Medical Journal Editors (ICMJE), *Defining the Role of Authors and Contributors, in Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals*, May 2023, https://www.icmje.org/recommendations/. [last accessed 15 September 2025].

[50] Multiple academic and professional bodies have affirmed that AI tools should not be listed as authors but may be acknowledged as part of the research process. The Committee on Publication Ethics (COPE) states that AI cannot assume accountability and thus does not meet the criteria for authorship. Similarly, the Association for Computing Machinery (ACM), the American Chemical Society (ACS), and the American Society of Civil Engineers (ASCE) all maintain that generative AI tools lack legal agency and intellectual responsibility and should be treated as non-authoring tools whose use must be transparently disclosed.

[51] Government Technology, *Can ChatGPT be credited as an author?*, 2023. https://www.govtech.com/question-of-the-day/can-chatgpt-be-credited-as-an-author. [last accessed 15 September 2025].

4.4 *Bias, Accuracy, and the Need for Human Judgment*

AI language models come with well-documented limitations that carry ethical implications for scholarly writing. They can reflect biases present in their training data, produce inaccurate information, and lack the ability to judge truth or relevance. Integrating such tools into academic work thus poses risks to the quality and objectivity of research literature.

One significant implication of using AI tools in scholarly writing is the potential amplification of existing societal biases. Suppose the training data of a language model reflects underrepresentation or stereotypes related to gender. In academic writing, such tendencies can subtly distort how research problems are framed or how populations are represented. A recent study in "Nature" cautioned that large language models may «amplify social biases like sexism and racism» if their outputs are not critically reviewed and corrected[52].

Ethically responsible use of AI requires scholars to remain vigilant for implicit biases in generated text. When AI-produced sections systematically omit certain literatures or perspectives – such as privileging Western sources over others – the human author must intervene to restore balance by integrating the missing viewpoints. Much like an editor correcting bias in a human-authored draft, the academic responsibility here lies with the author to edit and critically assess the AI-generated content.

Another critical concern in the use of large language models is their tendency to generate hallucinations and factual inaccuracies. These models do not possess inherent fact-checking capabilities and often produce incorrect information with a high degree of fluency and confidence. There have been prominent cases in which AI tools fabricated academic citations that, upon verification, were found to be entirely non-existent – an unacceptable error in scholarly contexts.

In fact, in early May 2025, "The New York Times" reported that an OpenAI's investigation into its latest GPT-o3 and GPT-o4-mini LLMs were found to hallucinate between 33% and 48% more compared to earlier versions[53].

As a result, human oversight remains indispensable. While AI tools may expedite the drafting process, the time saved must be reinvested in meticulous fact-checking and validation. Each AI-generated sentence should be cross-referenced with credible sources or verified calculations. This is not merely a matter of due diligence – it is a safeguard against the illusion of precision that AI-generated prose can create. Scholars must resist the temptation to accept fluent outputs at face value and instead apply their critical expertise to assess both the factual and logical soundness of the content. As highlighted in various institutional guidelines, responsible use of AI means treating its outputs as drafts requiring confirmation, not as definitive contributions[54]. The ultimate authority in scholarly communication

[52] Hu, T., Kyrychenko, Y., Rathje, S., et al., *Generative language models exhibit social identity biases*. "Nature", Vol.5 no.1, pp. 1-10, 2024. https://doi.org/10.1038/s43588-024-00741-1. [last accessed 15 September 2025].
[53] According to the New York Times, GPT- o3 hallucinated 33% of the time and GPT-o4 mini 48% of the time when tested on the PersonQA benchmark, significantly higher than the hallucination rate of OpenAI's earlier model GPT-o1. Metz, C., A.I. Is *Getting More Powerful, but Its Hallucinations Are Getting Worse*, The New York Times, 2025. https://www.nytimes.com/2025/05/05/technology/ai-hallucinations-chatgpt-google.html. [last accessed 15 September 2025].

must remain with the human author, whose judgment and analytical rigor cannot be delegated to a machine.

Another challenge posed by using generative AI in academic writing is the lack of source transparency. Most large language models do not cite their sources unless explicitly designed for that purpose – a feature absent in the majority of general-use models. Consequently, authors who rely on AI tools must assume the additional responsibility of sourcing any factual assertions produced by the system.

At the same time, when used critically, AI tools can also contribute to improving scholarly rigor. Through interactive prompts or dialogic exchanges, AI can identify weak points in an argument or suggest areas requiring clarification. Some researchers use tools like ChatGPT as a form of intelligent interlocutor – soliciting critiques, probing questions, or counterarguments in response to draft paragraphs[55]. This engagement can prompt authors to reflect more deeply, refine their reasoning, and address potential gaps in their logic. However, such benefits are contingent on the user's discernment. This type of supervised interaction suggests that, when guided by human expertise, AI can function as a tool to enhance analytical depth and thoroughness in academic work.

As it was already mentioned, international bodies emphasize the importance of human judgment in any AI deployment. In research, this translates to a simple rule: "AI can assist, but humans must decide". The scientist's discernment is the final safeguard to ensure that what goes into the scientific record is accurate and unbiased.

5. *Typology of AI Usage in Academic Writing*

Not all forms of AI assistance in academic writing are equivalent. It is important to differentiate levels of involvement, which may range from minimal linguistic support to the extensive generation of content. In this context, we propose a typology of AI use in scholarly writing, categorized according to the degree and nature of the system's intervention.

Level 0 – No AI Assistance: The manuscript is written entirely by the human author, aside from standard software like spelling/grammar checkers inherent in word processors (which we consider baseline tools). There is no contribution from generative AI and this level represents traditional writing and serves as a reference point.

[54] As emphasised in recent institutional frameworks, generative AI outputs should not be regarded as final academic contributions. Rather, they must be critically assessed, fact-checked, and revised by human researchers who remain fully accountable for the content. The European Commission's *Living Guidelines on the Responsible Use of Generative AI in Research* explicitly state that researchers must remain aware of the limitations of generative tools – including hallucinations, bias, and inaccuracies – and must treat all AI-assisted content as provisional. Similarly, UNESCO's *Guidance for Generative AI in Education and Research* underscores the need for human-centred governance and advocates for ethical validation and oversight in all uses of generative AI in academic contexts. European Commission, *Directorate-General for Research and Innovation*... cit.

[55] Liang, W., et al., *Can large language models provide useful feedback on research papers? A large-scale empirical analysis*, arXiv preprint arXiv:2310.01783, 2023. https://doi.org/10.48550/arXiv.2310.01783. Developed an automated pipeline using GPT-4 to provide comments on full PDFs of scientific papers. Their evaluation revealed that GPT-4's feedback overlapped significantly with human peer reviewers' comments, suggesting its utility in pre-review processes.

Level 1 – AI as Copyediting Aid: AI is used for "non-substantive editing" and polishing of text. This includes grammar correction, spell-checking, formatting references, or improving syntax and style at the sentence level. The content and ideas are all human-generated; the AI just refines expression. For example, an author might use a tool like Grammarly or an LLM-based rephraser to fix awkward sentences or to suggest more concise wording. This is akin to having a very advanced proofreader. Many journals explicitly permit such use without requiring formal disclosure, considering it similar to using any writing enhancement software. However, some guidelines recommend acknowledging even this level if the changes are significant or if a generative AI (like ChatGPT) was used for it, just to be transparent[56]. Ethically, Level 1 is low risk as long as the author reviews all changes to ensure meaning is preserved.

Level 2 – AI for Ideation and Outlining: AI is used to "generate ideas", "suggest outlines", or "summarize sources", but the actual prose of the manuscript is written by the human author. In this scenario, a researcher might prompt an AI for brainstorming («What are potential subtopics I should cover in my literature review on X?») or ask it to explain a concept as a starting point, or to condense a set of articles into key points. The AI might also help organize thoughts (e.g., proposing a logical section structure). However, when it comes to writing paragraphs and crafting arguments, the human does the heavy lifting. This may enhance both efficiency and creativity, functioning as a form of ideation support or conceptual scaffold. Since the AI's role is limited to the planning phase, it typically does not result in any directly AI-generated text appearing in the final manuscript. Some disclosure might still be prudent (e.g., «Note: an AI tool was used to assist in initial brainstorming for this article»), but since all final text is original to the author, the main ethical consideration is ensuring the ideas taken from AI are properly verified and cited if they are not common knowledge. Level 2 blurs into standard research work (researchers often get "ideas" from many sources), except that AI can provide those ideas quickly on demand.

Level 3 – AI-Generated Draft Passages with Human Revision: At this level, the AI generates some portions of text, which the human author then 2heavily edits", "fact-checks", and "integrates" into the manuscript. For instance, an author could have ChatGPT draft a generic introduction based on key points the author provides, then rewrite most of it in their own voice, adding specifics and correcting any AI errors. Or an author might use AI to create a rough summary of a certain subtopic, then refine it extensively. Here, the AI is a co-drafter, but the human remains the chief writer who curates and modifies the content. This mode of use requires meticulous oversight: the author must ensure no factual inaccuracies survive and that the style is made consistent. Disclosure is essential at this level, as portions of the text did originate from AI (even if altered). A statement might read: «An initial draft of the Background section was generated with the assistance of ChatGPT and was subsequently revised for accuracy and completeness by the authors». Some journals might ask for identifying which sections had AI input. The ethical acceptability of Level 3 depends on the degree of human revision. When revisions are merely superficial, the outcome may resemble the characteristics of Level 4 use. However,

[56] For instance, Elsevier permits the use of AI-assisted technologies to improve the readability and language of a manuscript. Authors are expected to maintain oversight and control, ensuring the content's accuracy. Disclosure is encouraged to promote transparency. The IEEE clarifies that while the use of AI tools for editing and grammar enhancement is common practice and generally does not require disclosure, any content generated by AI (text, figures, images, code) must be disclosed in the acknowledgments section. SAGE differentiates between assistive AI tools and generative AI. Tools that enhance language, grammar, or structure are considered assistive and do not require disclosure. However, the use of generative AI tools that produce content must be disclosed.

if the AI-generated output is treated as raw material and substantively reworked by the human author, its role may be more accurately described as that of an assistant or ghostwriter operating under the author's direction. This level constitutes a critical area of ongoing debate, as it occupies a grey zone: to what extent does AI-generated content compromise authorial integrity? How can we ensure that the author's intellectual contribution and voice remain central? In such cases, transparent disclosure and clear attribution of responsibility are essential safeguards to uphold academic accountability.

Level 4 – AI-Generated Text with Minimal Human Edits: In this category, significant parts of the manuscript are written by an AI and the human author does only light editing – fixing a few sentences or errors but largely accepting the AI's output. For example, an author might prompt an AI to "write a literature review on Y" and then copy-paste the result into their paper with only minor changes. This level is highly problematic and not considered responsible practice by most standards. The human's contribution to the writing is marginal; effectively the AI is doing the composition, raising questions of originality and authorship. Even if sources are accurate, the author cannot claim to have independently constructed the narrative or arguments. Such use is likely to violate academic standards and the editorial policies of many journals or conferences, even if disclosed[57]. If such an approach is taken, it would require at least clear and explicit disclosure – for example: «Section 2.1 was generated by ChatGPT using prompts provided by the author and has undergone only minor editing». However, it is questionable whether such a paper would be publishable in a reputable journal, as it fails the criterion of being the author's original work. Level 4 use may be more applicable in non-research contexts, such as the production of books or institutional reports, where authors might employ AI tools to generate large volumes of prose. In the context of academic research, however, this level of reliance on AI is widely regarded as unethical – comparable to outsourcing the writing of a manuscript and subsequently claiming authorship.

Level 5 – Fully AI-Generated with Human Curatorial Role: This extreme would be a paper essentially written by AI from start to finish, with the "author" only curating the process (providing prompts, maybe selecting the best outputs). The human might not do any substantial writing or editing at all. At Level 5, the human role is largely reduced to assembling AI-generated outputs or prompting the system through iterative refinements. Ethically, this level is not meaningfully distinct from Level 4; in fact, it may be more problematic, as the human contribution to the written content becomes minimal or even negligible. The resulting work amounts to automated writing, and the submission of such material as original scholarship would constitute a clear violation of authorship and academic integrity standards. Under existing academic norms, this is considered a form of academic misconduct, insofar as the human presents AI-generated content as their own.

The only context in which this use might be acceptable is within scholarship that explicitly examines AI-generated text as an object of study. In such cases, the inclusion of AI-produced content must be clearly labelled and confined to demonstrative purposes. However, as a general method for producing research papers, Level 5 lies outside the boundaries of acceptable academic practice.

[57] When authors use generative AI tools to draft portions of a manuscript – such as an initial introduction or summary – and then significantly revise the content for accuracy, coherence, and style, most academic publishers require transparent disclosure. Leading publishers like Elsevier, Springer Nature, and JAMA emphasize that even when the AI's contribution is heavily edited, its initial involvement must be acknowledged. These policies aim to maintain authorship integrity while allowing responsible integration of AI in the writing process.

While some have speculated about a future in which AI systems could autonomously produce publishable research, such a scenario would fundamentally challenge prevailing definitions of authorship and scholarly responsibility. Current policies address this concern pre-emptively by prohibiting AI from being credited as an author.

Most real-world cases of AI use in 2023/2024 fall into Levels 1–3. Levels 4–5 are discouraged and would likely lead to rejection or retraction if discovered. It is useful for institutions to define these levels because it helps communicate what is acceptable[58].

By delineating levels, we move away from a binary "AI or not" view, and toward a nuanced understanding that some AI uses are relatively low impact (similar to hiring an editor), whereas others fundamentally shift who is actually doing the intellectual work away from the researcher.

6. *Framework for AI Transparency: Toward an "AI Use Transparency Index"*

To operationalise the above guidelines and typology, we propose developing an "AI Use Transparency Index" (AUTI) as a framework for standardizing how researchers report AI involvement in their work. The goal of the AUTI is to promote consistent disclosure and set clear expectations for transparency. This index would rate or categorize the extent of AI usage and corresponding transparency measures. It is considered useful that an AI Use Transparency Index (AUTI) be promoted and established as a framework by an international body such as the Council of Europe or UNESCO, rather than as a voluntary element arising from the initiative of a few institutions or research groups dealing with these issues.

Key components of the AI Use Transparency Index might include:

Level of AI Intervention: A straightforward classification of how much AI contributed to the writing. For example, an author might be required to self-identify the usage level (1 through 5) when submitting a manuscript. This level could then be published alongside the article (e.g., a small note on the first page: «AI Use Transparency Index: Level 2 – AI-assisted brainstorming and language editing, with all content written by authors».). This gives readers immediate context about AI's role. It also holds authors accountable to not misrepresent their usage, since claiming a lower level than actually used would be an ethical breach.

Disclosure Statement Quality: The index would encourage not just "if" AI was disclosed, but "how thoroughly". It could have criteria or a scoring for disclosure statements. For instance, a complete disclosure might need to mention the specific tool (e.g., GPT-4), the version, the exact purpose (editing vs drafting vs summarizing), and an affirmation of human verification. An AUTI scoring rubric could award higher transparency scores for disclosures that include all these elements. In contrast, a vague

[58] *UNESCO Guidance for Generative AI in Education and Research* proposes key steps to the regulation of GenAI tool, including mandating the protection of data privacy, and setting an age limit for the independent conversations with GenAI platforms. To guide the proper use of the tools in education and research, the Guidance proposes a human-agent and age-appropriate approach to the ethical validation and pedagogical design processes.

statement like “an AI was used” without details would be considered low transparency. Journals could set a required threshold (e.g., a minimum transparency score) for acceptance.

Verification and Accountability Measures: Part of the framework may involve authors describing how they verified AI-generated content. This could be an additional statement or checklist. The AUTI could include an indicator (yes/no or brief description) of the verification process. For example: «All AI-generated text was reviewed for accuracy against primary sources and edited for clarity by the authors». Permission and Compliance Checks: The framework can tie into institutional compliance by logging whether the AI usage was in line with the relevant policies. For instance, some sensitive research might forbid using external AI due to data privacy. The AUTI disclosure could note, «Use of AI was reviewed and approved by the Institutional Review Board (IRB) or Data Protection Officer» if applicable or simply affirm that no confidential data was exposed to AI.

Standardization of Reporting Language: Over time, an AUTI would benefit from standard phrases or templates. Consistent phrasing can make it easier for meta-analyses or searches to identify how AI was used across the literature. For example, a standardized statement: «The authors used [Tool name] (version X) to assist with [specific task]. The authors have reviewed and take responsibility for the content generated».

Scoring System (Optional): One might envision an AI transparency “score” on a scale (say 0 to 5 or 0 to 10) for each paper, where a higher score means more transparent and responsible use reporting. For instance, a paper that used no AI or only minimal AI (Level 0–1) might score 5/5 on transparency by stating that clearly. A paper that used AI but gave a full account might also score 5/5. But a paper that used AI and gave an incomplete disclosure or needed editor extraction might score lower.

The proposed AI Use Transparency Index (AUTI) is designed to benefit multiple stakeholders in scholarly communication. For authors, it provides clear expectations on how to disclose AI use; for reviewers, it offers a standardized checklist to assess compliance; for readers, it ensures immediate visibility into a paper’s AI context; and for ethicists, it facilitates the tracking of emerging trends. Beyond its practical utility, AUTI would also serve a normative function by formalizing AI disclosure and reinforcing that non-disclosure is unacceptable.

7. *Conclusions and Recommendations*

The integration of artificial intelligence into academic writing presents both opportunities and challenges. While it offers unprecedented support in articulating ideas and streamlining the drafting process, it also raises critical ethical and practical concerns. This paper has examined key dimensions of this evolving landscape – including transparency, plagiarism, bias, oversight, authorship, and institutional policy – alongside the emerging responses from academic journals, universities, and international organizations committed to preserving the integrity of scholarly communication in the context of AI.

At its core, responsible AI use in academia implies the augmentation of scholarly work without

compromising its integrity. When deployed appropriately, AI can assist researchers in refining expression, reducing linguistic barriers, and enhancing productivity, particularly for those with limited proficiency in the language of publication or with accessibility needs. It may also serve as a catalyst for creativity by suggesting alternative formulations or perspectives. However, these potential benefits are contingent upon the author's sustained critical engagement with the writing process. AI should operate strictly as a supportive instrument – not as a substitute for genuine intellectual labour. Its use becomes ethically problematic, and may constitute academic misconduct, when it obscures authorship, circumvents scholarly rigor, or introduces unverified content without transparent disclosure.

As of 2025, there is a growing consensus – reflected in institutional policies and editorial guidelines – that AI-assisted writing is permissible only when accompanied by full transparency and retained human responsibility throughout the research and publication process.

The academic community is actively contributing to the development of normative frameworks to guide the ethical integration of AI into scholarly practice. Given the pace of technological advancement, sustained dialogue among researchers, educators, editors, and ethicists will be essential to refine standards and ensure that academic values are upheld in an increasingly AI-mediated research environment.

Considering the analysis, several actionable recommendations can guide the ethical and effective integration of AI in academic writing across stakeholder groups.

For Researchers (Authors): Authors should approach AI as a supportive tool – useful for refining expression or generating ideas – while retaining complete intellectual control and critical oversight. Any AI involvement must be disclosed with specificity in publications. Researchers are encouraged to stay informed about discipline-specific guidelines, consult ethical bodies when in doubt, and contribute to collective learning by sharing their experiences with AI use in scholarly work.

For Academic Journals and Publishers: Journals should establish and publicly display clear policies on AI authorship and usage. Submission systems should include mandatory AI disclosure statements, and editors and reviewers should be trained to assess compliance. Journals should harmonize standards, adopt frameworks like the AI Use Transparency Index, and ensure responsible use of AI-detection tools, always accompanied by human judgment.

For Universities and Educational Institutions: Academic integrity policies must be updated to explicitly address generative AI, distinguishing between permitted and prohibited uses. AI literacy training should be embedded across curricula, covering not only technical use but also ethical citation and disclosure. Educators should receive support to redesign assessments that emphasize critical thinking. For Research Funders and Policymakers: Funding agencies should incorporate AI disclosure requirements into grant conditions and support empirical research on the impact of AI on scholarly communication. Policymakers, including those shaping legislation like the EU AI Act, should consult with academic communities to balance innovation and integrity. International organizations such as UNESCO and OECD should continue facilitating dialogue and producing harmonized guidance, particularly to support lower-capacity institutions. Moreover, companies providing AI tools should be encouraged to implement transparency features aligned with academic needs.

For Students and Early-Career Researchers: As they form foundational research habits, students and junior scholars should be taught to use AI as a learning aid, not a substitute for their own work. Responsible experimentation – for instance, comparing AI outputs with their own writing – can foster skill development. Mentorship and open dialogue can help guide ethical usage. Early-career researchers should also be encouraged to engage publicly in conversations on AI use, contributing to the evolving academic norms with their insights and lived experiences.

Finally, the academic community has shown a notable degree of adaptability in response to the rapid emergence of generative AI technologies. Rather than hindering scholarly progress, the responsible integration of AI into academic writing practices may serve to enhance inclusivity, support multilingual researchers, and improve overall research efficiency – so long as its use is governed by transparency, critical oversight, and a steadfast commitment to academic integrity. Moving forward, it is essential that authors explicitly acknowledge their use of AI and retain full responsibility for the content produced. Under these conditions, generative AI need not be viewed as a threat, but rather as a constructive partner in the pursuit and dissemination of knowledge.

Bibliography

ALLEA, *The European Code of Conduct for Research Integrity*, 2023. *https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/european-code-of-conduct-for-research-integrity_horizon_en.pdf*. [last accessed 15 September 2025].

Bubeck, S., Chandrasekaran,V., Eldan, R., et al., *Sparks of Artificial General Intelligence: Early experiments with GPT-4*, arXiv preprint arXiv:2303.12712, 2023. https://doi.org/10.48550/arXiv.2303.12712.

CGTN, *University of Hong Kong issues interim ban on ChatGPT, AI-based tools*, 19 febbraio 2023.

https://news.cgtn.com/news/2023-02-19/University-of-Hong-Kong-issues-interim-ban-on-ChatGPT-AI-based-tools-1hxWzqgcMxy/index.html. [last accessed 15 September 2025].

Council of Europe (CoE), *Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*, "Council of Europe Tereaty Series", no. 225, 5 September 2024. https://rm.coe.int/1680afae3c. [last accessed 15 September 2025].

COPE, *Artificial intelligence and authorship*, 2023. https://publicationethics.org/news-opinion/artificial-intelligence-and-authorship. [last accessed 15 September 2025].

COPE, *Authorship and AI tools*, 2023. https://publicationethics.org/guidance/cope-position/authorship-and-ai-tools. [last accessed 15 September 2025].

Cyberspace Administration of China, *Interim Measures for the Administration of Generative Artificial Intelligence Services*, 2023.

Elsevier, *Generative AI policies for journals*, n.d. https://www.elsevier.com/about/policies-and-standards/generative-ai-policies-for-journals. [last accessed 15 September 2025].

Elsevier, *The use of generative AI and AI-assisted technologies in writing for Elsevier,* n.d. https://www.elsevier.com/about/policies-and-standards/the-use-of-generative-ai-and-ai-assisted-technologies-in-writing-for-elsevier. [last accessed 15 September 2025].

EU–US Trade and Technology Council (TTC), T*erminology and Taxonomy for Artificial Intelligence*, Annex to the Joint Roadmap on Evaluation and Measurement Tools for Trustworthy AI and Risk Management, European Commission & U.S. Government, 2024.

EU Artificial Intelligence Act, *The EU Artificial Intelligence Act Up-to-date developments and analyses of the EU AI Act.*, n. d. https://artificialintelligenceact.eu. [last accessed 15 September 2025].

European Commission (EC), *Artificial Intelligence Act. Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence*, Brussels, European Commission, 2024.

European Journal of Therapeutics, *Authorship and artificial intelligence (AI) tools*, Gaziantep University Faculty of Medicine, 2025. https://eurjther.com/index.php/home/Authorship-AI. [last accessed 15 September 2025].

European Commission, Directorate-General for Research and Innovation, *Responsible Use of Generative AI in Research. Living Guidelines on the Responsible Use of Generative AI in Research*, ERA Forum Stakeholders' Document, Second Version, April 2025, Brussels, European Commission, 2025.

European Commission, *Living guidelines on the responsible use of generative AI in research*. 15 April 2025. https://research-and-innovation.ec.europa.eu/document/2b6cf7e5-36ac-41cb-aab5-0d32050143dc_en. [last accessed 15 September 2025].

Government Technology, *Can ChatGPT be credited as an author?*, 2023. https://www.govtech.com/question-of-the-day/can-chatgpt-be-credited-as-an-author. [last accessed 15 September 2025].

Harvard University, *Research with Generative AI*, n.d. https://www.harvard.edu/ai/research-resources/. [last accessed 15 September 2025].

Hosseini, M., Resnik, D. B., Holmes, K., *The ethics of disclosing the use of artificial intelligence tools in writing scholarly manuscripts*, "Research Ethics", Vol.19, no. 4, 2023. https://doi.org/10.1177/17470161231180449.

Hu, T., Kyrychenko, Y., Rathje, S., et al., *Generative language models exhibit social identity biases*. "Nature" Vol. 5, n.1, 2024, pp. 1-10. https://doi.org/10.1038/s43588-024-00741-1.

International Committee of Medical Journal Editors (ICMJE), *Defining the Role of Authors and Contributors, in Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals*, May 2023, https://www.icmje.org/recommendations/. [last accessed 15 September 2025].

Lau, A., *The stance of academic journals on the use of AI*, Elion, 2023. https://elion.nz/the-stance-of-academic-journals-on-the-use-of-ai/. [last accessed 15 September 2025].

Liang, W., et al., *Can large language models provide useful feedback on research papers? A large-scale empirical analysis*, arXiv preprint arXiv:2310.01783, 2023. https://doi.org/10.48550/arXiv.2310.01783.

Ministry of Education, Singapore, *AI in Education (AIEd) Ethics Framework*, "Student Learning Space", Singapore, 2025. https://www.learning.moe.edu.sg/ai-in-sls/responsible-ai/ai-in-education-ethics-framework/. [last accessed 15 September 2025].

Meckler, L., *New York City schools ban AI chatbot ChatGPT for fear of cheating and misinformation*, "The Washington Post", 2023. https://www.washingtonpost.com/education/2023/01/05/nyc-schools-ban-chatgpt/. [last accessed 15 September 2025].

Metz, C., A.I. *Is Getting More Powerful, but Its Hallucinations Are Getting Worse*, The New York Times, 2025. https://www.nytimes.com/2025/05/05/technology/ai-hallucinations-chatgpt-google.html. [last accessed 15 September 2025].

Nature, *Tools such as ChatGPT threaten transparent science; here are our ground rules for their use*, 2023. https://www.nature.com/articles/d41586-023-00191-1. [last accessed 15 September 2025].

New York University Steinhardt, *Academic integrity and syllabus support in the age of generative AI,* NYU Steinhardt, 2023. https://steinhardt.nyu.edu/faculty-and-staff/academic-affairs/steinhardt-ai-hub/academic-integrity-and-syllabus-support-age. [last accessed 15 September 2025].

Nature, *Tools such as ChatGPT threaten transparent science; here are our ground rules for their use*, 2023. https://www.nature.com/articles/d41586-023-00191-1. [last accessed 15 September 2025].

OECD, *Explanatory Memorandum on the OECD Framework for the Classification of AI Systems and the Updated OECD Definition of an AI System*, Paris, OECD, 2024.

OECD, *Recommendation of the Council on Artificial Intelligence (OECD AI Principles)*, OECD Legal No. 0449, Paris, OECD, 2024. https://oecd.ai/en/ai-principles. [last accessed 15 September 2025].

Singapore Management University, *Disclosure on the use of AI in research manuscripts: How are researchers doing it?*, 2024. https://library.smu.edu.sg/topics-insights/disclosure-use-ai-research-manuscripts-how-are-researchers-doing-it. [last accessed 15 September 2025].

Stanford Institute for Human-Centered Artificial Intelligence, *Researchers use GPT-4 to generate feedback on scientific manuscripts*, 2023. https://hai.stanford.edu/news/researchers-use-gpt-4-generate-feedback-scientific-manuscripts [last accessed 15 September 2025].

Stokel-Walker, C., *ChatGPT listed as author on research papers: Many scientists disapprove*, "Nature", 2023 https://www.nature.com/articles/d41586-023-00107-z. [last access 15 September 2025].

ELife *Publishing and Peer Review at eLife*, n.d. https://elifesciences.org/about/peer-review. [last accessed 15 September 2025].

Taylor & Francis, *Taylor & Francis Clarifies the Responsible Use of AI Tools in Academic Content Creation*, Taylor & Francis Newsroom, 17 February 2023. https://newsroom.taylorandfrancisgroup.com/taylor-francis-clarifies-the-responsible-use-of-ai-tools-in-academic-content-creation/. [last accessed 15 September 2025].

The Guardian, *Science journals ban listing of ChatGPT as co-author on papers*, "The Guardian", 2023. https://www.theguardian.com/science/2023/jan/26/science-journals-ban-listing-of-chatgpt-as-co-author-on-papers. [last access 15 September 2025].

The AI Pact encourages and supports organisations to plan ahead for the implementation of AI Act measures. European Commission, *AI Pact*. https://digital-strategy.ec.europa.eu/en/policies/ai-pact. [last access 15 September 2025].

Thorp, H. H., *ChatGPT is fun, but not an author*, "Science", Vol. 379, no. 6630, 2023. https://doi.org/10.1126/science.adg787.

Toner, H., *What are generative AI, large language models, and foundation models?*, Center for Security and Emerging Technology (CSET), 2023. https://cset.georgetown.edu/article/what-are-generative-ai-large-language-models-and-foundation-models/. [last access 15 September 2025].

UNESCO, *Recommendation on the Ethics of Artificial Intelligence*, Paris, UNESCO, 2021, disponibile su: https://unesdoc.unesco.org/ark:/48223/pf0000381137. [last access 15 September 2025].

UNESCO, *Guidance for Generative AI in Education and Research*, Paris, UNESCO, 2023, disponibile su: https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research. [last access 15 September 2025].

Vincent, J., *ChatGPT can't be credited as an author, says world's largest academic publisher*, The Verge, 2023. https://www.theverge.com/2023/1/26/23570967/chatgpt-author-scientific-papers-springer-nature-ban. [last access 15 September 2025].