https://doi.org/10.17323/jle.2025.29876

# The 2025 Landscape of Generative AI in Scholarly Writing and Publishing: A Scoping Review of Uses and Ethical Approaches

**Lilia Raitskaya** [1], **Elena Tikhonova** [2]

[1] Moscow State Institute of International Relations (MGIMO University), Moscow, Russia
[2] Peoples' Friendship University of Russia (RUDN University), Moscow, Russia

## ABSTRACT

**Introduction:** The rapid advancement of generative artificial intelligence (GenAI) has outpaced earlier reviews of its role in scholarly writing. Scholarship is shifting from problem-framing to explicitly normative work emphasising transparency, accountability, and sustained human oversight, yet the operationalisation of ethical guidance in editorial and authorial practice remains insufficiently systematised.

**Purpose:** This scoping review maps 2025 evidence on AI applications in academic publishing and identifies emerging normative frameworks that enable workflow efficiencies while preserving human intellectual ownership and accountability.

**Method:** Using the Arksey and O'Malley framework and PRISMA-ScR reporting, we systematically searched Scopus for English-language articles and reviews published in 2025. Eligibility criteria were defined via the PCC framework. Included publications were charted and analysed thematically to capture use cases, governance responses, and ethical concerns.

**Results:** The search identified 334 records, with 56 publications meeting the inclusion criteria. The corpus shows global authorship and, after manual verification, an approximately balanced mix of reviews and primary studies, revealing substantial document-type misclassification in the database. Discourse clusters around governance (authorship and policy), technological impact (content quality), and risk mitigation (academic integrity). Prominent use cases include support for intellectual tasks (ideation, outlining, and synthesis), language enhancement, and support in peer review and editorial workflows; each catalyses distinct ethical challenges. In response, structured normative frameworks, such as tiered disclosure models and task-based AI taxonomies (e.g., GAIDeT), are emerging to make accountability auditable while preserving human oversight. Across the sample, AI is positioned as an assistive tool subordinate to human responsibility; immediate ethical regulation dominates, whereas educational integration and broader cultural critique remain secondary. We outline a research agenda focused on framework validation, improved detection infrastructures, longitudinal cognitive outcomes, human–AI collaboration design, policy standardisation, and decolonial analyses of algorithmic bias.

**Conclusion:** The field is moving from problem identification toward solution-oriented governance. Progress now depends on interdisciplinary efforts that translate normative principles into workable publishing procedures, ensuring GenAI strengthens, rather than undermines, academic integrity and equitable knowledge production.

## KEYWORDS

generative artificial intelligence; academic publishing; research integrity; authorship governance; disclosure models; human oversight; AI task taxonomy; GAIDeT; policy standardisation

**Citation:** Raitskaya, L., & Tikhonova, E. (2025). The 2025 Landscape of Generative AI in Scholarly Writing and Publishing: A Scoping Review of Uses and Ethical Approache. *Journal of Language and Education, 11*(4), 5-50. https://doi.org/10.17323/jle.2025.29876

**Correspondence:**
Elena Tikhonova,
etihonova@hse.ru

**Received:** December 08, 2025
**Accepted:** December 15, 2025
**Published:** December 30, 2025

## INTRODUCTION

The rapid integration of generative artificial intelligence (GenAI) into the academic landscape has catalysed one of the most significant shifts in scholarly communication in decades, sparking a complex and urgent ethical debate. A primary focus of this discourse has been on establishing robust norms for transparency, requiring clear disclosure of AI use, and upholding the non-negotiable principle of human oversight and authorial responsibility (Cheng et al., 2025). This conversation intensified dramatically following the public release of ChatGPT 3.5, which initially led to publications controversially listing the AI as a co-author (King & ChatGPT, 2023). This practice, however, was quickly met with scrutiny, unleashing a broader academic conversation that moved beyond simple authorship attribution to critically examine the nature, limits, and ethical implications of AI's contribution to scholarly work (Hufton, 2023; Quasem, 2023; Knöchel et al., 2025).

In response to this ferment, a consensus has begun to coalesce around a tiered model of ethical use, which prioritises applications that augment rather than replace human intellect. This framework categorises AI use from the ethically sound (such as refining grammar, improving readability, and translating text with human verification) to the ethically contingent, including generating outlines or brainstorming ideas, but only when based on substantial human-provided input and critical steering (Cheng et al., 2025). The core of this model is the principle of human primacy, which asserts that the primary ideas, critical analysis, intellectual substance, and final accountability for a manuscript must originate from and be owned by human authors. Consequently, leading editorial bodies like the Committee on Publication Ethics (COPE, 2024)[1] and the International Committee of Medical Journal Editors (ICMJE, 2025)[2] have explicitly prohibited attributing authorship to AI tools, arguing that they cannot assume accountability for the work. Their use, however, can and should be formally acknowledged in the manuscript (Yoo, 2025).

The top-down guidance from the publishing ethics committees has created an urgent and pragmatic need for publishers, journals, research institutions, and universities worldwide to translate these principles into concrete, actionable guidelines. The absence of clear, institution-specific policies creates a vacuum filled by ambiguity and inconsistent practices, potentially compromising academic integrity and creating ethical risks (Alberth, 2023). Publishers require standardised protocols for manuscript screening and reporting; journals need clear instructions for authors and reviews on permissible AI use; universities and research institutions must develop training, ethical frameworks, and authorship policies that address the use of these tools in grant writing, student work, and research dissemination. The current period is this characterised by a race to adapt, with stakeholders across the academic ecosystem grappling with how to properly address the current stance on AI appliances in a way that is both principled and practical (Lim B. et al., 2025).

The COPE guidelines (2024) underscore this complexity, emphasising that authors are not only responsible for the validity and originality of their content but also for ensuring that any use of AI does not breach copyright or introduce biased on fabricated information. COPE recommends that the use of AI be thoroughly and transparently described in the methods section or another appropriate part of the manuscript. Such comprehensive disclosure is no longer a mere recommendation but is increasingly seen as a cornerstone of contemporary research integrity, essential for maintaining trust and allowing for the critical appraisal of the research process by editors, reviewers, and readers (Hosseini et al., 2025).

Concurrent with this policy evolution, a burgeoning body of research continues to refine ethical practices, with a marked annual increase in relevant publications. This literature closely tracks the evolving use patterns of AI and AI-powered tools across the entire scholarly writing lifecycle (Bozkurt, 2024; Bobier et al., 2025). A significant focus is on the educational imperatives for researchers, ensuring they develop the necessary digital and ethical competencies for responsible engagement. Simultaneously, there is a pronounced emphasis on delineating ethically sound applications of AI at every stage of the scientific publishing workflow (from the initial literature review and data analysis to manuscript drafting, peer review, and the verification of references) aiming to integrate technological innovation seamlessly with unwavering academic integrity (Erdat et al., 2025).

The decision to undertake this new systematic scoping review in 2025, following our earlier works in 2023 and 2024 (Tikhonova & Raitskaya, 2023; Raitskaya & Tikhonova, 2024), is a direct response to the unprecedented velocity of change characterizing the field of generative AI. The knowledge landscape in this domain is not merely expanding. It is fundamentally transforming at a pace that renders even recent syntheses quickly outdated. Our earlier works captured the initial shockwave of ChatGPT's release and the emergent themes of ethical panic and potential. However, the subsequent period has seen the rapid maturation of AI tools, the development of more nuanced and varied policies from publishers and institutions, and a significant body of new empirical research on real-world implementation and outcomes.

[1] COPE. Authorship and AI Tools. https://publicationethics.org/guidance/cope-position/authorship-and-ai-tools

[2] ICMJE Recommendations. https://icmje.org/recommendations/

This review is therefore necessary to map this newly evolved territory, capturing the consolidation of certain practices, the resolution of early debates, and the emergence of new, more complex challenges that were only nascent a year ago. While the reviews by Khalifa & Albadawy (2024), Raitskaya & Tikhonova (2024), Fabiano et al. (2024), Yan et al. (2024) and others effectively mapped the initial landscape of AI applications and challenges in academic writing, they were based on literature from the immediate post-ChatGPT release period (2023-2024). A 2025 review is critical to capture the subsequent wave of empirical studies, institutional policy implementations, and evolving ethical frameworks that have emerged as the technology has become more deeply integrated into scholarly workflows.

This new review will provide an updated and more comprehensive synthesis, moving beyond initial explorations to document established use-cases, longitudinal impacts, and the effectiveness of proposed mitigation strategies for ethical challenges. Furthermore, it addresses a specific gap left by previous publications by systematically investigating how high-level ethical guidelines are being operationalized by researchers, publishers, and institutions. By focusing explicitly on the intersection of "Uses" and "Ethical Approaches," the 2025 review aims to offer a crucial, forward-looking resource to guide the responsible, effective, and equitable integration of generative AI into the very foundation of knowledge production and dissemination.

The review aims to synthesise and summarize uses of AI and AI-powered appliances in scholarly writing and publishing in 2025.

The research questions:

RQ 1. What are the basic features and distribution of publications examining AI uses in scholarly writing and publishing?
RQ 2. What are the predominant thematic clusters that characterise the scholarly discourse on the use of AI and AI-powered tools in academic writing and publishing?
RQ 3: How do prevailing patterns of AI application in scholarly writing shape the associated ethical discourse, and what normative frameworks are emerging in response?
RQ 4. What are the future research agendas for the field?

# METHOD

## Protocol

To ensure methodological rigor, a comprehensive research protocol was developed prior to commencing this scoping review. This final report constitutes a transparent and accurate representation of the review as conducted. The methodological approach was informed by the stages set forth by Arksey and O'Malley (2005) and reported in accordance with the PRISMA-ScR guidelines (Tricco et al., 2018). The review was carried out in strict adherence to the pre-established protocol, with no recorded departures.

## Eligibility Criteria

The problem, concept, and context (PCC) were defined to establish an effective search strategy (see Table 1), with a rationale for each criterion. A discussion and some calibration exercises were performed before the criteria were finally established.

## Search Strategy

The selection of keywords was an iterative process, informed by the review objective, research questions, and a scoping of the current literature. These candidate terms were evaluated and refined through a series of preliminary searches to finalize the most effective terminology. The subsequent definitive search was conducted in the Scopus database using the following keyword strings:

(1) (AI) AND (scholarly writing OR academic writing OR research reporting OR AI-assisted writing OR research workflow OR manuscript preparation);
(2) (ethics) AND (research) AND (AI-assisted writing).

The literature search for this scoping review was executed concurrently by both authors on the same date to maintain procedural uniformity. The initial pool of identified publications was then subjected to the eligibility process, with inclusion contingent upon relevance and full-text availability.

## Study Selection

The study selection process was conducted as follows:

(1) Initial identification: both reviewers independently identified publications using the pre-defined keywords and eligibility criteria.
(2) Database filtering: preliminary screening was performed using Scopus internal filters for publication data, language, subject area, and document type.
(3) Title and abstract screening: the reviewers independently screened the titles and abstracts of the filtered results, tagging each record as "include" or "exclude". Any discrepancies were resolved through discussion until a consensus was reached.
(4) Full-text retrieval: the full texts of provisionally included studies were sourced from publishers' websites or by directly contacting the corresponding authors via academic networks.
(5) Final eligibility assessment: both reviewers then performed a full-text review of the retrieved documents independently to make the final determination on eligibility.

**Table 1**

*Eligibility Criteria*

| Criterion | Inclusion | Exclusion | Rationale |
|---|---|---|---|
| Problem | Uses of AI and AI-powered applianc-es in scholarly writing and publishing | All publica-tions beyond the theme | The problem encompasses the entire workflow from literature search and data analysis to manuscript drafting and peer review |
| Concept | Scholarly writing and ethics of scholarly publishing | Other concepts | The concepts cover the ways AI is used and its implications for (1) Scholarly Writing (quality, originality, authorial voice, transparency, and the very nature of academic authorship); (2) Ethics of Publishing (authorship attribution, dis-closure requirements, plagiarism, reproducibility, bias in AI systems, and the integrity of the scientific record) |
| Context | Research at large and academic journals | Other contexts | The context clearly situates the review within the specific environment where these issues are most relevant.<br>Higher Education and Research: Captures the perspective of researchers, fac-ulty, and graduate students who are primary users of these tools.<br>Academic Journals: Captures the perspective of publishers, editors, and peer reviewers who are gatekeepers of the scholarly record and are creating poli-cies in response to AI |
| Language | English | Other languages | The object of all research in focus is scholarly publications in English. The lan-guage choice is also identified by its status as a lingua franca of international science |
| Time period | All publications for 2025 | Not applica-ble | The review aims to synthesise the current knowledge on the AI uses in re-search writing and publishing |
| Types of sources | Full texts of articles, reviews, and edito-rials indexed at the Scopus database | Unavailable sources, un-available full texts | The chosen types of publications represent the most authoritative and foun-dational pillars of the scientific literature. Full-text articles (particularly primary research) provide the original data, methodologies, and empirical evidence. Review articles offer a critical synthesis of this primary literature, distilling findings from numerous studies to map the current state of knowledge. Ed-itorials are a crucial scholarly genre as they interpret the significance of new discoveries and often identify nascent fields, ethical dilemmas, and paradigm shifts before they are fully evident in the primary literature |
| Geogra-phical location | Any location | Not applicable | The choice was deliberately designed to encompass a global scope to capture the worldwide dissemination, adoption, and discourse surrounding, avoiding the limitation of a single regional context |
| Database | Scopus | Other bases | Scopus was selected as the primary database for this scoping review due to its extensive and interdisciplinary coverage of high-quality, peer-reviewed litera-ture. As one of the largest curated abstract and citation databases, it provides robust indexing across the key subject areas central to this investigation al-lowing for a comprehensive mapping of the scholarly discourse. Its advanced search functionality and reliable export features facilitate a systematic and reproducible search strategy, ensuring that the review captures a represent-ative body of published research on the practices and ethical implications of Generative AI across the academic landscape. |
| Areas of Research | Social Sciences; Computer Science; Medicine; Arts and Humanities; Engi-neering; Decision Sciences | Other areas | The selection of subject areas for this review (Social Sciences, Computer Science, Medicine, Arts and Humanities, Engineering, and Decision Sciences) is strategically designed to capture the multifaceted discourse on Genera-tive AI in academia. This interdisciplinary approach is necessary because the conversation is siloed across domains that each provide an indispensable perspective: Computer Science and Engineering literature forms the techni-cal foundation, detailing the capabilities and limitations of the tools them-selves; Medicine and Social Sciences offer critical insights into the practical application and behavioural impact of these tools in high-stakes, empirical research environments; the Arts and Humanities contribute the foundational philosophical and critical analysis of core ethical concepts like authorship, originality, and creativity that are being fundamentally challenged; and final-ly, Decision Sciences provides the crucial governance perspective, capturing literature on institutional policies, ethical frameworks, and strategic man-agement of AI integration. Without this comprehensive scope, the review would risk presenting a fragmented or incomplete map of the interconnected practices and ethical implications defining the field |

## Data Charting

A data-charting form was developed collaboratively by the authors. To calibrate the process, both reviewers independently piloted the form by extracting data from five included studies. The form was then refined through an iterative discussion. The final captured data variables included title of study, author(s), year of publications, type of publications, theme of publications, author keywords, AI or AI-powered appliances, uses of AI and AI-powered appliances, population if applicable, research design if applicable, type of review if applicable, key findings, future research agendas. All extracted data were subsequently verified by both authors to ensure accuracy.

## Data Analysis and Validation

The analysis of the charted data-utilised a mixed-methods approach, integrating descriptive synthesis with thematic analysis to address the research questions. The process began with a within-case analysis of individual studies, examining key contextual variables including research design, population, application of Generative AI, principal findings, and proposed research agendas. Subsequently, a cross-case analysis was conducted to systematically identify, code, and compare patterns, points of coverage, and divergences across the entire corpus of publications. This coding was performed independently by both reviewers and subsequently refined through iterative discussions to establish analytical consensus and reliability.

Emergent thematic clusters were further validated through a bibliometric co-occurrence analysis of keywords using VOSviewer. This integration of quantitative keywords mapping served to triangulate and substantiate the inductively derived themes with empirical trends from the literature itself. The final synthesis is therefore structured around these validated thematic clusters, each substantiated with representative evidence from the included studies. These clusters form the foundational framework for answering the research questions and for delineating the review conclusions regarding ethical directions, prevailing uses of AI scholarship, and consolidated agenda for future research.

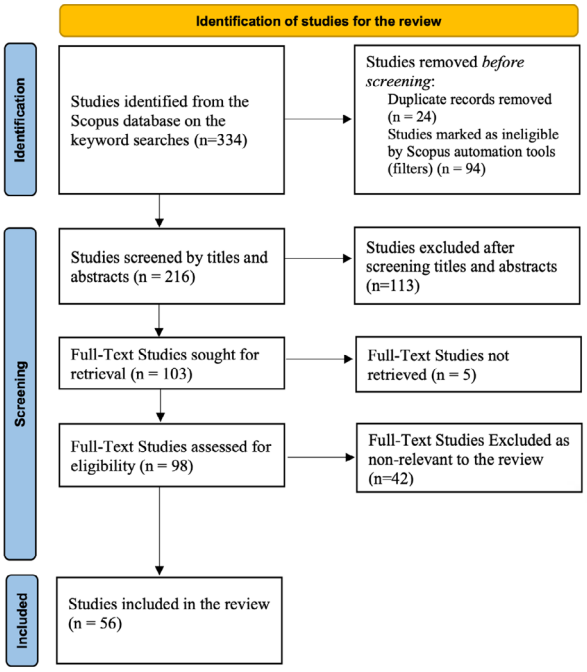
# RESULTS

## Search and Selection Results

A systematic search of the Scopus database was conducted on October 29, 2025, using predefined search strings related to AI and ethics in scholarly writing. The initial search yielded 334 records after removing 24 duplicates and applying database filters. Subsequent title and abstract screening reduced this number to 216. Following the retrieval and eligibility assessment of 103 full-text publications, 42 were excluded as irrelevant. Consequently, a final sample of 56 publications was included in the review. The PRISMA flowchart (Figure 1) depicts the whole identification and screening procedure.

**Figure 1**

*PRISMA 2020 Flow Chart. Selection of Publications for the Review*



## Publication Landscape: Volume, Distribution, and Key Characteristics

### *The Review Sample*

The final sample of 56 publications featured authors affiliated with 39 countries, led by the USA (n=12), India, and Turkey (n=6 each). A worldwide distribution of affiliations is shown in Figure 2. The most prolific institutions, such as Hong Kong Polytechnic University and Northwestern University Feinberg School of Medicine, contributed two publications each, while the vast majority (44 organizations) were represented by a single publication. Similarly, the author Hosseini, M. was the most frequent contributor (n=2), with 158 authors appearing only once. The average number of authors per publication was 2.9. *Discover Education* was the leading journal (n=3). Six other journals each published two articles: *Accountability in Research*, *Computers and Education: Artificial Intelligence*, *Forum for Linguistic Studies*, *Humanities and Social Sciences Communications*, *Journal of Korean Medical Science*, and *Proceedings of the ACM on Human-Computer Interaction*. The remaining 49 journals contributed one publication each.

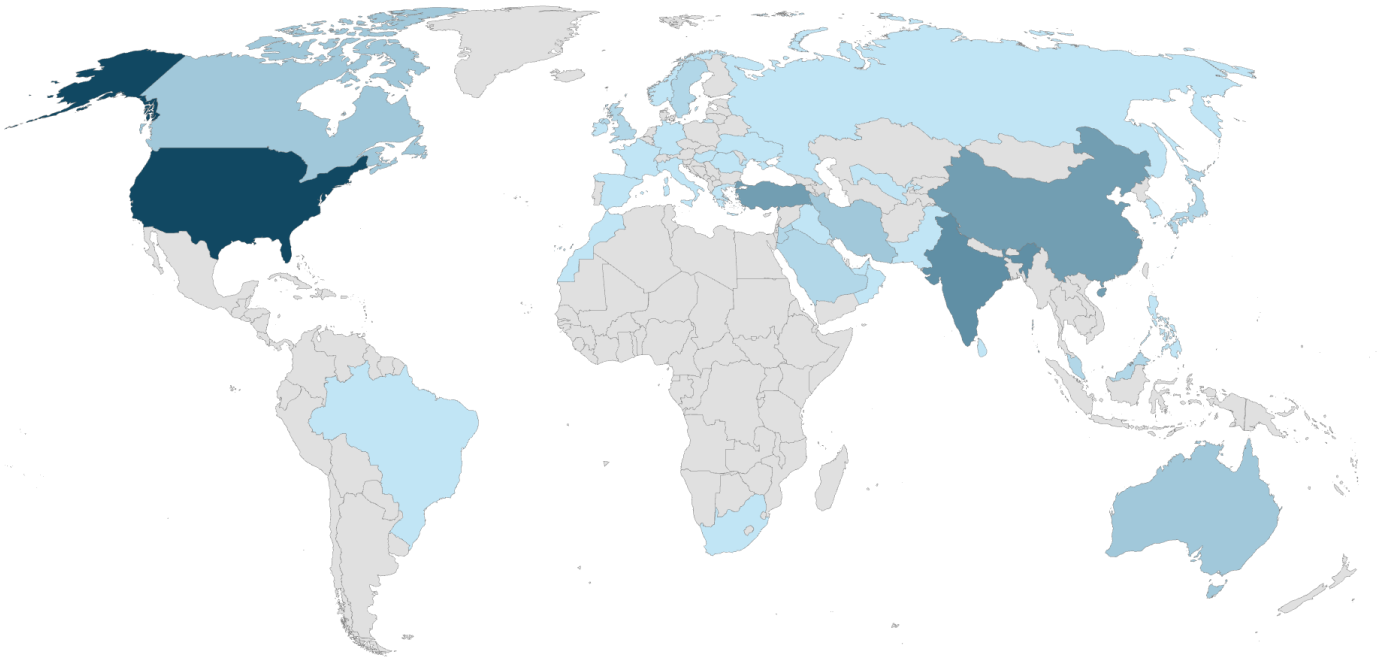
The analysis of the 56 publications included in this review reveals a significant discrepancy between the formal document type assigned by the Scopus database (Figure 3) and the actual methodological nature of the work (Figure 4). While Scopus provides a standardized classification, a manual content analysis demonstrates that these labels can be misleading, necessitating a deeper examination to understand the true composition of the scholarly landscape on this topic. The core of this issue lies in the fact that Scopus's automated or author-submitted classifications do not always accurately reflect the research design, leading to a substantial number of reviews being mis-categorised as articles.

This misclassification is quantitatively significant. According to the provided appendices, while only eight publications are formally grouped as "reviews" (Appendix 1), a further ten publications listed as "articles" in Appendix 2 are, upon inspection, explicitly described as reviews of various types. For instance, Alamri et al. (2025) is marked as an "article" in Scopus but methodologically is a "Systematic Review combined with a SWOT analysis." Similarly, Knight (2025) and Liu et al. (2025) are both identified as "articles" in the database, yet their research designs are a "Scoping Review" and a "Systematic Scoping Review," respectively. Other clear examples include Alnaimat et al. (2025) (a narrative/ perspective review), Anglhelescu et al. (2025) (a scoping review), and Lendvai (2025) (a scientometric review). This indicates that nearly half of the reviews in the total sample (10 out of 18 total reviews) are incorrectly labelled by the database, which could skew bibliometric analyses that rely solely on Scopus's document type filter.

The implications of these discrepancies are considerable for researchers relying on database searches. A scholar search-

**Figure 2**

*A Worldwide Breakdown of Affiliations of Authors' Publications Included in the Review*



*Note. The legend includes the density of publications per country, with a maximum of 12 (the darkest shade) to a minimum of 1 (the lightest shade). The map was made up in the Microsoft excel by the authors*

ing specifically for reviews on this topic in Scopus would miss over half of the relevant synthetic literature. Conversely, a search filtered for "Article" would be contaminated with numerous reviews, potentially comprising the purity of a sample intended for analysing primary empirical studies. This highlights the critical importance of manual, content-based verification in systematic and scoping reviews, as automated database classifications are insufficiently reliable. The data from Appendices 1 and 2 collectively show that the actual body of literature consists of 1 editorial, 18 reviews (8 formal and 10mis-categorised), and 37 primary articles or non-review conceptual papers, a breakdown that is fundamentally different from the initial Scopus-derived count of 47 articles and 8 reviews (Figures 3 and 4).

Furthermore, despite the initial classificatory discrepancies, the final, verified sample is remarkably well-balanced, incorporating a diverse spectrum of publication types that significantly strengthens the ensuing synthesis. The inclusion of a foundational corpus of 18 reviews, comprehensive analyses like the hybrid systematic review by Arar et al. (2025) and the scoping review by Ghasemi et al. (2025), provides a robust, high-level overview of the field's themes and debates. This is powerfully complemented by the 37 primary articles and conceptual papers detailed in Appendix 2, which offer granular, empirical insights and original perspectives. This combination creates a synergistic effect: the reviews map the landscape, while the articles allow for a detailed examination of the individual studies. This methodological triangulation ensures that the resulting synthesis is not only comprehensive but also grounded, capable of capturing both the established consensus from aggregated reviews and the cutting-edge, often contested, knowledge emerging from primary research.

A deeper examination of the 47 publications initially classified as articles reveals a rich spectre of research designs, underscoring the multi-faceted approach required to study in academic writing. The sample includes a strong contingent of empirical studies employing diverse methodologies. These range from experimental and quasi-experimental designs (e.g., Akgun et al., 2025; Kumar et al., 2025; Erol et al., 2025) that test hypotheses and compare AI versus human output, to observational and cross-sectional studies (e.g., Alkhawan et al., 2025; Archana et al., 2025; Nassar et al., 2025) that document prevalence, attitudes, and correlations. Furthermore, qualitative and mixed-methods approaches are well-represented, providing nuanced insights into human experiences and processes. Studies like Hu et al. (2025) and Picardal et al. (2025) use interviews and descriptive analysis to explore identity negotiation and pedagogical integration, while Comas-Forgas et al. (2025) and Hwang et al. (2025) combine quantitative and qualitative data for a more complete picture. This methodological pluralism is complemented by non-empirical contributions, including conceptual analyses (Cambraia & Pyrrho, 2025), policy reviews (Bobier et al., 2025), and philosophical essays (Hytten, 2025), which are crucial for developing the ethical and theoretical frameworks guiding the field.

Similarly, the 18 reviews in the sample demonstrate a sophisticated application of synthesis methodologies beyond a monolithic "review" label. The corpus includes systematically conducted reviews that follow rigorous protocols for searching and selecting evidence, such as the Systematic Review with a SWOT framework by Alamri et al. (2025) and the Scoping Reviews by Ghasemi et al. (2025) and Raitskaya & Tikhonova (2025). These are complemented by bibliometric analyses (Goyibova et al., 2025; Lendvai, 2025), which quantitatively map the intellectual structure and evolution of the research field. On the more narrative end of the spectrum, there are perspective and narrative reviews (Mondal et al., 2025; Ateriya et al., 2025) that provide and synthesise literature without a formal systematic methodology. Notable, the sample also features hybrid models, such as the work by Arar et al. (2025), which combines bibliometric performance analysis with thematic content analysis.

This diversity in review types is a significant strength, as it allows for the synthesis of knowledge through different lenses. They include mapping the field's metrics, systematically aggregating findings on specific questions, and providing critical, expert-led overviews of the complex ethical and practical landscape.

The review primarily focuses on Generative AI Large Language Models (LLMs), with ChatGPT being the most extensively studied tool. Other types include detectors, paraphrasers, and specialized research assistants. AI uses for image or code generation, or their analyses are less researched within the theme. The table below (Table 2) sums up the breakdown of the specific tools studied, mentioned, or otherwise touched upon in the 56 documents of the review.

## Core Themes in the Literature on AI-Powered Writing and Publishing

Following the coding process, the reviewed literature was synthesized into ten core thematic clusters, capturing the central topics of scholarly debate. Table 3 provides a comprehensive overview of these clusters and lists the publications associated with each, illustrating the scope and focus of current research. The coding of the publications was based on the themes of the publications (Appendix 5) and the consolidated key findings of the publications (Appendix 3).

### *VOSviewer Thematic Clustering of Keywords vs the Hypothesised Clusters*

The keyword co-occurrence analysis (with keyword density = 3) generated five thematically coherent clusters that illuminate how the scholarly community conceptualizes the uses of generative AI within research writing and academic pub-

**Figure 3**

*Types of Publications Included in the Review as Marked in the Scopus Database*

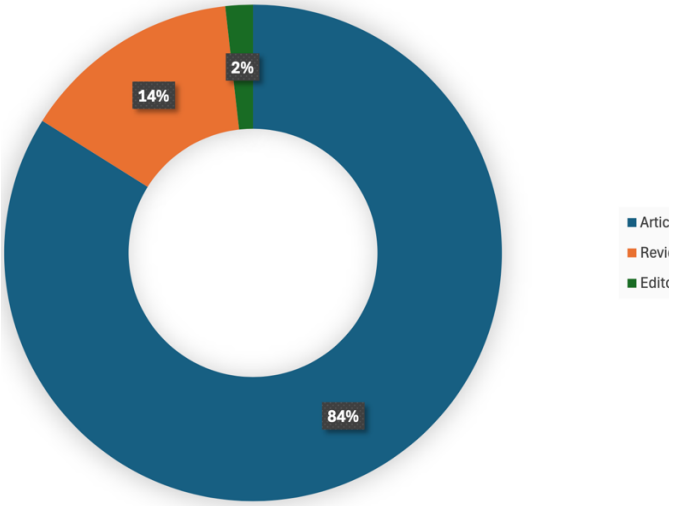



**Figure 4**

*Types of Publications Included in the Review by Content Analysis*

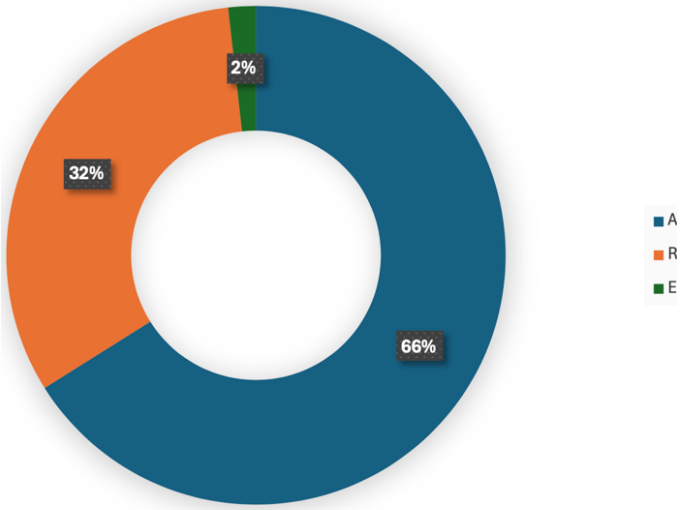



**Table 2**

*AI and AI-Powered Tools Studied in the Publications Included in the Review*

| Type of AI Tool | Specific Tools Mentioned | References |
|---|---|---|
| **Generative AI / Large Language Models (LLMs)** | ChatGPT (all versions, especially 3.5, 4), Google Gemini/Bard, Anthropic's Claude, Microsoft Copilot/Bing AI, Meta AI (Llama), ERNIE Bot, Perplexity AI, DeepSeek | Akgun et al., 2025; Al Hosni, 2025; Alamri et al., 2025; Alghazo et al., 2025; Alnaimat et al., 2025; Anghelescu et al., 2025; Arar et al., 2025; Archana et al., 2025; Ateriya et al., 2025; Bobier et al., 2025; Cambraia & Pyrrho, 2025; Cohen & Moher, 2025; Comas-Forgas et al., 2025; Daoudi, 2025; Ebadi et al., 2025; Erel et al., 2025; Erol et al., 2025; Fakharifar et al., 2025; Fung & Ng, 2025; Ghasemi et al., 2025; Goyibova et al., 2025; Hosseini et al., 2025; Hu et al., 2025; Huang et al., 2025; Hwang et al., 2025; Hytten, 2025; Kumar et al., 2025; Lendvai, 2025; Lim G.H. et al., 2025; Liu et al., 2025; Miller et al., 2025; Mo & Crosthweite, 2025; Mohan et al., 2025; Mondal et al., 2025; Moorhouse et al., 2025; Nassar et al., 2025; Öztürk et al., 2025; Parlak et al., 2025; Picardal et al., 2025; Prakash et al., 2025; Premat et al., 2025; Pretorius et al., 2025; Radtke & Rummel, 2025; Rafi & Amjad, 2025; Ragel, 2025; Raitskaya & Tikhonova, 2025; Resnik & Hosseini, 2025; Reza et al., 2025; Sparkman & Witt, 2025; Suchikova et al., 2025; Veiga et al., 2025; Yaqin et al., 2025; Yin & Chapelle, 2025; Yoo, 2025 |
| **AI Detection Tools** | ZeroGPT, GPTZero, Copyleaks, Originality.AI, Corrector App, iThenticate (Turnitin), Scribbr AI Detector | Alnaimat et al., 2025; Erel et al., 2025; Erol et al., 2025; Miller et al., 2025; Mondal et al., 2025; Nassar et al., 2025; Yoo, 2025 |
| **AI Paraphrasers/ "Humanizers"** | QuillBot, Stealth Writer, Grammarly (partly) | Alamri et al., 2025; Goyibova et al., 2025; Mondal et al., 2025; Picardal et al., 2025; Pretorius et al., 2025 |
| **AI-Powered Research & Writing Assistants** | Grammarly, Quillbot, Scispace, ChatPDF, Jenni AI, Typeset, Scite, Mendeley, EndNote, Zotero, Open AI's Deep Research, Scopus AI, CAQDAS (MAXQDA, NVivo), Google NotebookLM | Alamri et al., 2025; Archana et al., 2025; Alnaimat et al., 2025; Goyibova et al., 2025; Moorhouse et al., 2025; Picardal et al., 2025; Ragel, 2025; Reza et al., 2025; Veiga et al., 2025 |
| **AI for Image/ Code Generation & Analysis** | Midjourney, DALL-E, Sora, Leonardo AI, GitHub Copilot, OpenAI Codex, Proofig, ImageRights, TinEye | Alnaimat et al., 2025; Fung & Ng, 2025; Pretorius et al., 2025; Yin & Chapelle, 2025 |

**Table 3**

*Distribution of Publications among Hypothesised Thematic Clusters*

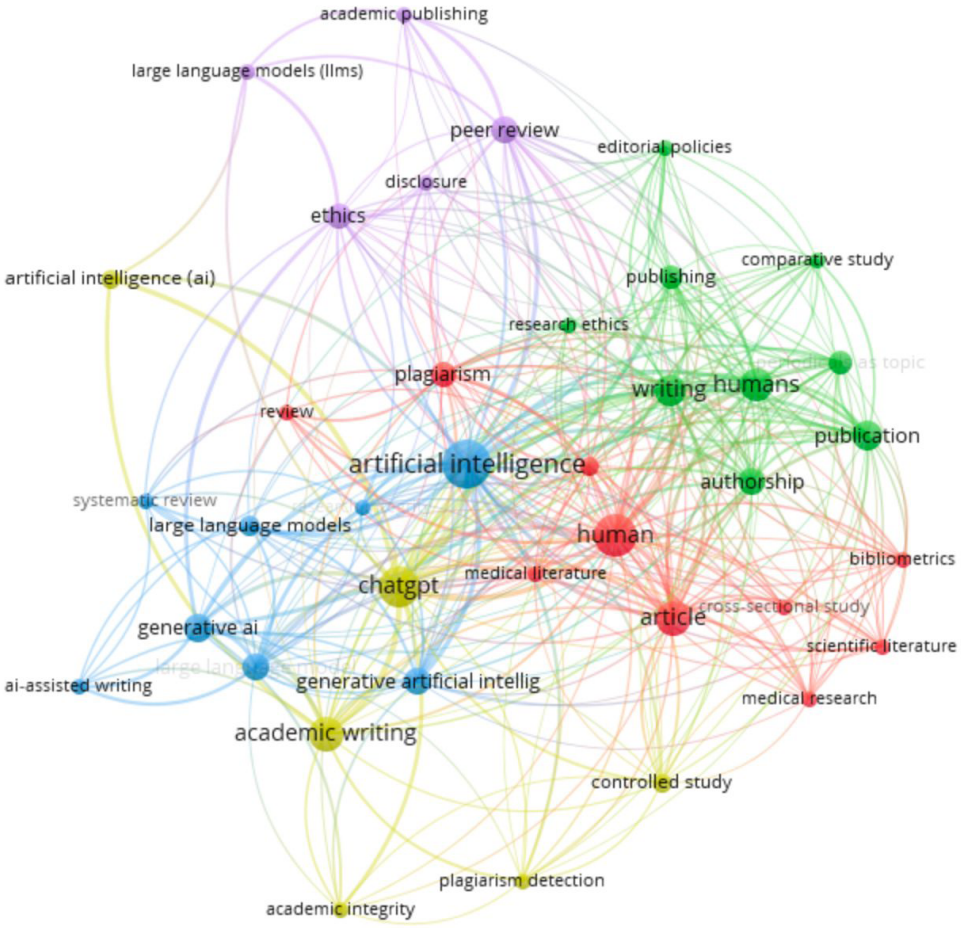
| Nos | Thematic Clusters | Included Publications |
|---|---|---|
| 1. | Journal Policies and Editorial Guidelines on AI Use | Alkhawam et al., 2025; Alnaimat et al., 2025; Bobier et al., 2025; Cohen & Moher, 2025; Fakharifar et al., 2025; Hosseini et al., 2025; Huang et al., 2025; Knight, 2025; Lim G.H. et al., 2025; Mohan et al., 2025; Moorhouse et al., 2025; Resnik & Hosseini, 2025; Veiga et al., 2025; Yin & Chapelle, 2025; Yoo, 2025 |
| 2. | Ethical Issues, Authorship, and AI Use Disclosure | Akgun et al., 2025; Al Hosni, 2025; Arar et al., 2025; Ateriya et al., 2025; Cambraia & Pyrrho, 2025; Daoudi, 2025; Erel et al., 2025; Goyibova et al., 2025; Hytten, 2025; Liu et al., 2025; Miller et al., 2025; Mondal et al., 2025; Nassar et al., 2025; Öztürk et al., 2025; Parlak et al., 2025; Premat et al., 2025; Resnik & Hosseini, 2025; Suchikova et al., 2025; Veiga et al., 2025; Yoo, 2025 |
| 3. | Quality, Detection, and Reliability of AI-Generated Content | Akgun et al., 2025; Anghelescu et al., 2025; Erel et al., 2025; Erol et al., 2025; Hwang et al., 2025; Miller et al., 2025; Mo & Crosthweite, 2025; Mondal et al., 2025; Nassar et al., 2025; Öztürk et al., 2025; Parlak et al., 2025; Pretorius et al., 2025; Yaqin et al., 2025 |
| 4. | Impact of AI on Academic Writing and Learner/Researcher Skills | Al Hosni, 2025; Alamri et al., 2025; Alghazo et al., 2025; Archana et al., 2025; Comas-Forgas et al., 2025; Goyibova et al., 2025; Hu et al., 2025; Hwang et al., 2025; Kumar et al., 2025; Liu et al., 2025; Mo & Crosthweite, 2025; Picardal et al., 2025; Prakash et al., 2025; Radtke & Rummel, 2025; Rafi & Amjad, 2025; Raitskaya & Tikhonova, 2025; Reza et al., 2025; Sparkman & Witt, 2025; Yaqin et al., 2025 |
| 5. | AI Applications and Potentials in Specific Disciplines and Research Types | Akgun et al., 2025; Anghelescu et al., 2025; Cohen & Moher, 2025; Ebadi et al., 2025; Fakharifar et al., 2025; Fung & Ng, 2025; Ghasemi et al., 2025; Kumar et al., 2025; Lim H.G. et al., 2025; Miller et al., 2025; Mohan et al., 2025; Nassar et al., 2025; Parlak et al., 2025; Prakash et al., 2025; Ragel, 2025; Sparkman & Witt, 2025 |
| 6. | Educational Aspects, Integration, and AI Literacy Development | Al Hosni, 2025; Archana et al., 2025; Hu et al., 2025; Knight, 2025; Kumar et al., 2025; Moorhouse et al., 2025; Picardal et al., 2025; Premat et al., 2025; Pretorius et al., 2025; Radtke & Rummel, 2025; Rafi & Amjad, 2025; Raitskaya & Tikhonova, 2025; Reza et al., 2025; Yin & Chapelle, 2025 |
| 7. | Peer Review and Editorial Processes Involving AI | Alnaimat et al., 2025; Ebadi et al., 2025; Hosseini et al., 2025; Lim G.H. et al., 2025; Mohan et al., 2025; Öztürk et al., 2025 |
| 8. | Bibliometric and Scientometric Analysis of AI Publications | Arar et al., 2025; Goyibova et al., 2025; Lendvai, 2025 |
| 9. | Human-AI Collaboration, Authorial Voice, and Agency | Al Hosni, 2025; Hu et al., 2025; Hwang et al., 2025; Hytten, 2025; Mondal et al., 2025; Pretorius et al., 2025; Radtke & Rummel, 2025; Rafi & Amjad, 2025; Reza et al., 2025; Suchikova et al., 2025 |
| 10. | Social, Cultural, and Philosophical Aspects of AI Integration in Academia | Al Hosni, 2025; Cambraia & Pyrrho, 2025; Hu et al., 2025; Hytten, 2025; Prakash et al., 2025; Premat et al., 2025; Pretorius et al., 2025; Raitskaya & Tikhonova, 2025 |

lishing. These clusters do not merely enumerate lexical patterns; rather, they reflect underlying epistemic orientations, normative tensions, and evolving methodological practices that define the current state of the field. Below, each cluster is interpreted with attention to its conceptual scope and contribution to the broader discourse (Figure 5).

**Cluster 1. Foundational Scholarly Communication and Traditional Research Norms.** Cluster 1 (keywords: article, bibliometrics, cross-sectional study, human, medical literature, medical research, plagiarism, practice guideline, review, scientific literature) delineates the baseline conceptual landscape into which debates on generative AI are introduced. The terminology foregrounds core genres of scholarly communication (article, review, scientific literature) and standard methodological descriptors (cross-sectional study, medical research), revealing that much of the early discourse is anchored in biomedical and clinical research contexts. The recurrent presence of plagiarism indicates that AI-generated texts initially entered academic discussions through the lens of misconduct and quality assurance. The co-presence of bibliometrics and practice guideline suggests that concerns extend beyond writing quality to include the downstream effects on citation patterns, evidence synthesis, and clinical protocol adherence. In this sense, Cluster 1 functions as a conceptual substrate that frames generative AI not as a disruptive entity but as an element interfacing with deeply institutionalized norms of scientific communication.

**Figure 5**

*VOSviewer Thematic Clustering of Keywords*




**Cluster 2. Authorship, Editorial Governance, and Normative Regulation.** Cluster 2 (keywords: authorship, comparative study, editorial policies, humans, periodicals as topic, publication, publishing, research ethics, writing) encapsulates the ethical–administrative response to the diffusion of generative AI across academic writing practices. The prominence of authorship and editorial policies reflects the rapid reorientation of journals and publishers toward explicit governance mechanisms clarifying acceptable and unacceptable uses of generative models. Terms such as research ethics and humans highlight a fundamental tension between machine-generated content and expectations of human intellectual accountability. The presence of comparative study points to a growing body of work systematically evaluating policy variation across publishers, disciplines, and national contexts. Rather than treating AI merely as a technical tool, Cluster 2 reveals the field's shift toward interrogating responsibility, disclosure, and the redefinition of scholarly credit, making this cluster central for understanding institutional adaptation.

**Cluster 3. Generative Technologies, Writing Support, and Epistemic Integrity.** Cluster 3 (keywords: AI-assisted writing, artificial intelligence, generative AI, generative artificial intelligence, large language model, large language models, research integrity, systematic review) forms the technological and epistemic core of the literature. It comprises direct references to generative systems (LLMs, AI-assisted writing) linked with research integrity, underscoring that technological capability is consistently evaluated through the prism of epistemic reliability. The co-location of systematic review indicates that the field is rapidly maturing and has begun generating higher-order evidence synthesis on AI's roles, risks, and benefits. This cluster thereby captures the dual framing of generative AI: as an enabler of scholarly productivity and as a potential disruptor of methodological rigor. Its thematic cohesion demonstrates that questions of validity, accuracy, hallucination, and verifiability dominate the scholarly agenda at the intersection of AI and knowledge production.

**Cluster 4. Academic Integrity, Detection Strategies, and Evaluative Methodologies.** Cluster 4 *(keywords: academic integrity, academic writing, artificial intelligence, ChatGPT, controlled study, plagiarism detection)* represents the empirical-evaluative segment of the discourse, where generative AI is operationalized as an object of controlled experimentation. Research in this cluster frequently assesses the detectability, stylistic signatures, and fidelity of ChatGPT-generated texts. The co-occurrence of *controlled study* and *plagiarism detection* suggests a methodological emphasis on validation trials (testing existing detection tools, developing new classifiers, and comparing AI-produced content with human writing). *Academic integrity* functions here as a conceptual anchor, indicating that these empirical investigations are framed not in isolation but within broader pedagogical and institutional concerns. Cluster 4 thus reflects a phase in the field's evolution dominated by questions of risk mitigation, verification, and the operational boundaries of authenticity.

**Cluster 5. Transparency, Disclosure, and Editorial Decision-Making Ecosystems.** Cluster 5 *(keywords: academic publishing, disclosure, ethics, large language model, peer review)* occupies the policy-oriented and meta-scientific frontier of the literature. Here, generative AI is situated within the broader ecosystem of academic publishing, particularly regarding expectations of transparency (*disclosure*) and ethical practice. The association of *peer review* with *large language model* highlights new lines of inquiry into AI-assisted reviewing, editorial triage, and decision support systems. This cluster signals a shift from reactive policy formulation to more proactive considerations of how AI can be responsibly integrated into editorial workflows. Its focus on ethics and governance positions Cluster 5 as the domain through which the academic community negotiates the long-term institutional settlement of generative AI in scientific communication.

## A Multilayered Structure of the Scholarly Discourse

Across clusters, a coherent trajectory becomes apparent. The field begins with traditional concerns about research norms and misconduct (Cluster 1), transitions to formal governance and ethical deliberation (Cluster 2), and consolidates a technologically grounded understanding of LLM capabilities and risks (Cluster 3). It then extends into empirical validation studies addressing integrity challenges (Cluster 4) and culminates in systemic reflections on the future of editorial practice (Cluster 5). Collectively, the clusters illustrate a multi-tiered ecosystem in which generative AI is simultaneously a writing aid, a methodological challenge, an ethical concern, and a catalyst for reconfiguring scholarly publishing infrastructures.

The comparison of the hypothesised and VOSviewer clusters (Table 4) reveals a strong overall alignment between the hypothesised thematic structure and the objective map of the discourse generated by VOSviewer, while also highlighting key refinements in how the field is conceptually organised.

### *Substantial Confirmation and Thematic Consolidation*

Several of the hypothesised clusters found a direct and clear reflection in the VOSviewer results. This is particularly true for themes related to governance and policy (1, 2), the technological core and integrity (3, 9), and editorial processes (7). VOSviewer confirms that these are the central, well-defined pillars of the emerging field, around which scholarly activity is most concentrated.

### *Thematic Merging and Re-Grouping*

The most significant structural difference is that VOSviewer produced a smaller number of larger, more consolidated clusters (5 vs. 10). This indicates that within the actual discourse some of the themes are intrinsically intertwined. For instance, the hypothesised clusters on policies (1), ethics (2), and peer review (7) were algorithmically merged into two broader, overlapping VOSviewer clusters (Clusters 2 and 5) that focus on the overarching systems of nominative regulation and the publishing ecosystem.

### *Shift in Emphasis from Pedagogy to Control*

A crucial insight is that educational aspects (6) and social and philosophical dimensions (10), which were hypothesised as distinct areas based on the keywords and themes of the publications, di not form their own clusters in the VOSviewer analysis. Instead, they were embedded within clusters concerned with academic integrity and detection (4) and epistemic integrity (3). This suggests that at the current stage, the academic community's focus is more sharply on the challenges and risks AI poses to established systems of quality control, knowledge validity, and scholarly integrity, rather than on proactive educational models or broad cultural critique.

### *The Objective Discourse Map*

While the hypothesis correctly identified the full spectrum of relevant topics, the VOSviewer analysis reveals their objective hierarchy and interconnections. The algorithm shows that the discourse is structured around a few core problem complexes, including (1) institutional response (policies, ethics); (2) technological engagement (LLMs, writing support); and (3) risk mitigation (detection, integrity). Our hypotheses 1-3, 7, and 9 accurately target these core areas.

Thus, the comparison demonstrates the high validity of our initial thematic framework. The objective data from VOSviewer provides refinement of the key areas defined in the hypotheses, their relative weight and internal relationships, showing that the current discourse is more heavily focused on the issues of governance, control, and techno-

**Table 4**

*Aligning Hypothesised and VOSviewer Clusters*

| Hypothesised Thematic Clusters | Corresponding VOSviewer Clusters | Rationale for Alignment |
|---|---|---|
| 1. Journal Policies and Editorial Guidelines on AI Use | Cluster 2<br>Authorship, Editorial Governance, and Normative Regulation<br>Cluster 5<br>Transparency, Disclosure, and Editorial Decision-Making Ecosystems | Direct Match<br>The focus on developing policies, authorship rules, and disclosure requirements from your hypothesis is central to the VOSviewer clusters dealing with normative regulation and editorial practice |
| 2. Ethical Issues, Authorship, and AI Use Disclosure | Cluster 2<br>Authorship, Editorial Governance, and Normative Regulation<br>Cluster 5<br>Transparency, Disclosure, and Editorial Decision-Making Ecosystems | Direct & Complete Match<br>Ethical concerns, authorship debates, and disclosure discussions are the core themes uniting these two VOSviewer clusters |
| 3. Quality, Detection, and Reliability of AI-Generated Content | Cluster 4<br>Academic Integrity, Detection Strategies, and Evaluative Methodologies | Direct Match<br>The issues of quality, detection, and reliability are directly reflected in the empirical studies on detection tools and content evaluation that define VOSviewer's Cluster 4 |
| 4. Impact of AI on Academic Writing and Learner/Researcher Skills | Cluster 1<br>Foundational Scholarly Communication and Traditional Research Norms<br>Cluster 4<br>Academic Integrity, Detection Strategies, and Evaluative Methodologies | Partial Match<br>The impact on skills is discussed within the context of traditional research norms (Cluster 1) and, more specifically, through the lens of academic integrity and assessment (Cluster 4) |
| 5. AI Applications and Potentials in Specific Disciplines and Research Types | Cluster 1<br>Foundational Scholarly Communication and Traditional Research Norms | Partial Match<br>Many discipline-specific studies (e.g., in medicine) from your hypothesis fall into the VOSviewer cluster that describes the established research fields being disrupted by AI |
| 6. Educational Aspects, Integration, and AI Literacy Development | Cluster 4<br>Academic Integrity, Detection Strategies, and Evaluative Methodologies | Indirect Match<br>In the VOSviewer data, the educational theme did not form a distinct cluster but was subsumed under the broader context of academic integrity, suggesting it is currently perceived through a lens of risk and control |
| 7. Peer Review and Editorial Processes Involving AI | Cluster 5<br>Transparency, Disclosure, and Editorial Decision-Making Ecosystems | Direct Match<br>The theme of peer review is explicitly identified as a key component of Cluster 5, which focuses on the publishing ecosystem |
| 8. Bibliometric and Scientometric Analysis of AI Publications | Cluster 1<br>Foundational Scholarly Communication and Traditional Research Norms | Direct Match<br>The term «bibliometrics» is explicitly present in Cluster 1, which encompasses meta-research on scientific communication itself |
| 9. Human-AI Collaboration, Authorial Voice, and Agency | Cluster 3<br>Generative Technologies, Writing Support, and Epistemic Integrity | Direct Match<br>The concepts of collaboration, writing support, and human agency are central to Cluster 3, which focuses on the technology's role in knowledge production |
| 10. Social, Cultural, and Philosophical Aspects of AI Integration in Academia | Cluster 3<br>Generative Technologies, Writing Support, and Epistemic Integrity<br>Cluster 5<br>Transparency, Disclosure, and Editorial Decision-Making Ecosystems | Indirect Match<br>These broad questions did not form a separate cluster in VOSviewer but are diffused into discussions about knowledge integrity (Cluster 3) and the ethics of publishing systems (Cluster 5) |

logical integration, while educational and broader cultural aspects are currently peripheral. These aspects are essentially seen through the lens of the more focused issues.

## Prevailing AI Uses, Ethical Discourse, and Emerging Normative Frameworks

The analysis of the sampled literature reveals a complex ethical landscape structured around seven interconnected themes, with a clear scholarly consensus forming around several core principles for responsible AI use in academia. The use of AI for substantive intellectual tasks (e.g., hypothesis generation, manuscript drafting, and data analysis) has provoked fundamental questions about authorship and accountability (Resnik & Hosseini, 2025; Yoo, 2025). Conversely, the development of AI for language enhancement and editorial support has generated complex debates around transparency and potential bias in scholarly evaluation, particularly concerning non-native English speakers (Hosseini et al., 2025; Yin & Chapelle, 2025). Furthermore, applications involving data processing and peer review assistance have raised critical issues regarding privacy, confidentiality, and the preservation of human judgement in core scholarly processes (Ebadi et al., 2025; Ateriya et al., 2025; Lim G.H. et al., 2025). This direct relationship between application patterns and ethical challenges underscores that the normative frameworks emerging in response, such as tiered disclosure model and usage taxonomies, are precisely targeted at governing specific use cases rather than AI implementation in abstract terms (Suchikova et al., 2025; Veiga, 2025).

The most fundamental theme concerns the reconceptualization of authorship, attribution, and accountability. A strong, cross-disciplinary consensus holds that AI tools cannot qualify as authors, a policy universally upheld by major editorial bodies like the International Committee of Medical Journal Editors (ICMJE, 2023), the Committee on Publications Ethics (COPE, 2023), and the World Association of Medical Editors (WAME) and is enforced by nearly all academic journals (Yoo, 2025; Bombier et al., 2025; Huang et al., 2025; Yin & Chapelle, 2025). The rationale, as explored by Veiga (2025) and Suchikova et al. (2025), is that AI lacks the legal identity and capacity for accountability required for authorship. This leads to a core principle articulated by Resnik & Hosseini (2025) and Fakharifar et al. (2025): human authors must retain full and ultimate responsibility for the entire content of a manuscript. This necessitates rigorous human oversight, wherein authors must control, fast-check, and meticulously edit all AI-generated content to ensure its accuracy, integrity, and freedom from plagiarism or factual "hallucinations" (Anghelescu et al., 2025; Akgun et al., 2025).

This principle of accountability is tightly linked to the challenge of transparency and disclosure. The literature reveals significant debate and policy fragmentation regarding implementation. Hosseini et al. (2025) highlight a key tension between the need for openness and the risk of introducing bias against non-native English speakers who use AI for language policing, a finding corroborated by the analysis of applied linguistics journals by Yin & Chapelle (2025), who note that less than half provide specific GenAI guidelines. In response, refined frameworks are being proposed. Resnik & Hosseini (2025) advocate for a tiered model where disclosure is mandatory only for "substantial" use, such as when AI generates hypotheses, writes manuscript sections, or analyses data. For supportive tasks, disclosure is optional, and for routine tasks, it is unnecessary This structured approach is echoed in the GAIDeT taxonomy by Suchikova et al. (2025) and is reflected in the flexible policies of journals like the Journal of Korean Medical Science (Yoo, 2025).

The ease of text generation has profound implications for academic integrity and plagiarism, giving rise to what Al Hosni (2025) and Goyibova et al. (2025) term "algiarism", implying the presentation of AI-generated text as one's own. This challenge is compounded by the documented unreliability of AI-detection tools, whish Enrol et al. (2025) and Mondal et al. (2025) show can produce false positives and easily circumvented. To uphold integrity, the community is establishing clear boundaries, with Yin & Chapelle (2025) and Veiga (2025) reporting that an overwhelming majority of journals prohibit citing AI-generated content as an authoritative source.

A further, deeply held concern is the potential erosion of scholarly rigour and critical thinking. Raitskaya & Tikhonova (2025) and Alamri et al. (2025) warn that over-reliance on AI could delegate essential academic skills, including deep analytical reasoning. Furthermore, Al Hosni (2025) and Hytten (2025) argue that AI-generated texts often lack a distinctive "authorial voice", leading to a homogenisation of academic prose. The ethical imperative, therefore, is to ensure AI serves as an aid to, not a replacement for intellectual labour (Hytten, 2025).

On a systematic level, the issues of algorithmic bias and inequity present significant challenges. Ateriya wt al. (2025) and Cohen & Moher (2025) demonstrate that AI tools can perpetuate biases embedded in their training data. Furthermore, Cambraia & Pyrrho (2025) frame the unequal access to advanced AI tools as "technological colonialism", a dynamic that Prakash et al. (2025) and Pretorius et al. (2025) argue risks consolidating knowledge production and marginalizing scholars from under-resourced institutions. The corresponding imperatives are to audit AI outputs for bias and promote equitable access to diverse tools (Prakash et al., 2025).

The use of AI also introduces serious data privacy and confidentiality risks. Ateriya et al. (2025) and Ebadi et al. (2025) identify the input of sensitive data into cloud-based AI platforms as a major threat. This has led to clear prohibitions, particularly against using public AI tools to handle confiden-

tial peer-review documents, as stated by journals like JAMA and emphasised by Ebadi et al. (2025).

Finally, the integrity of the peer review process itself is under ethical scrutiny. While AI could increase efficiency, Ebadi et al. (2025) and Lim G.H. et al. (2025) warn of confidentiality breaches and the loss of nuanced human judgement. The prevailing consensus, as seen in the policies analysed by Yoo (2025) and Bombier et al. (2025), is that human oversight remains irreplaceable, and the use of AI for analysing confidential manuscripts is strictly prohibited.

In synthesis, the discourse has moved from identifying problems to proposing structured solutions. The current landscape is characterised by evolving norms that strive to balance the promise of AI, as noted by researchers like Fakharifar et al. (2025) and Mohammadi et al. (2025), with the unwavering preservation of academic integrity, accountability, and equity as foundational values.

The discourse tends to discuss practical solutions related to more systematised and framework-based guidelines on AI uses in academic writing and publishing. The review publications entail two taxonomies. GAIDeT (Generative AI Delegation Taxonomy) (Suchikova et al., 2025) is a classification system that provides a structured framework for researchers to delegate specific tasks to Generative AI (like ChatGPT) throughout the research and publishing process. It clearly defines which tasks can be assigned to AI (e.g., literature search, code generation, editing) at both broad (macro) and detailed (micro) levels, while ensuring that human researchers retain full accountability and oversight for the final work. Its goal is to standardize and bring transparency to AI use in academia, without granting AI authorship. The article by Resnik & Hosseini (2025) moves beyond the *whether* of AI disclosure to tackle the more practical question of *when* it is ethically required. In response to conflicting policies from publishers, the authors develop a clear, principled framework arguing that disclosure should be mandatory only when the use of AI is both intentional and substantial. The primary goal is to make disclosure practices useful for credit, accountability, and reproducibility, while avoiding the burden of trivial disclosures.

We thoroughly analysed the frameworks by Suchikova et al. (2025) and by Resnik & Hosseini (2025) to merge the main points into a toolkit for AI use in scholarly writing and publishing (Appendix 6). By using this structured toolkit, the academic community can navigate the complexities of AI-assisted scholarship, harnessing its power for innovation while steadfastly upholding the core values of accountability, integrity, and intellectual rigor. This toolkit is our response to the needs of all stakeholders in the academia related to scientific publishing (researchers, editors, reviewers, journals, publishers, faculty and educators).

## Future Research Agendas

Based on the synthesis of the 56 publications included in the review, the following consolidated research agenda has been identified based on the extracted research agendas (Appendix 3), addressing the most pressing future directions for the study of GenAI in academic writing and publishing:

(1) Development and validation of ethical guidelines and frameworks. There is a strong consensus on the need to develop, refine, and validate comprehensive ethical guidelines for GenAI use in scholarly writing. Future research should focus on creating internationally recognised, interdisciplinary ethical standards that address transparency, accountability, and human oversight (Arar et al., 2025; Knight, 2025; Resnik & Hosseini, 2025). This includes context-specific applications (e.g., in education, healthcare, and non-English academic contexts) and evaluating their practical implementation and adherence (Premat et al., 2025; Rafi & Amjad, 2025).

(2) Enhancing detection and integrity tools. The limitations of current AI-detection tools (reliability issues, bias, and susceptibility to evasion) highlight the need for more robust and transparent detection mechanisms. Future work should aim to improve the accuracy of AI-text identification, reduce false positive or negative issues, and develop tools capable of identifying AI-generated images, references, and data (Erol et al., 2025; Alnaimat et al., 2025; Yoo, 2025). Research is also needed to assess the impact of AI-humanisers and paraphrasing tools on academic integrity (Mondal et al., 2025).

(3) Longitudinal and impact studies. A significant gap exists in understanding the long-term effects of GenAI on critical thinking, writing skills, authorial voice, and scholarly development. Longitudinal studies are recommended to track how AI tools influence learning outcomes, cognitive offloading, and the development of research identity over time (Kumar et al., 2025; Hu et al., 2025; Raitskaya & Tikhonova, 2025). This also includes examining the impact of AI on research equity, especially for non-native English speakers and underrepresented groups (Pretorius et al., 2025; Prakash et al., 2025).

(4) Human-AI collaboration and user-centred design. Research should explore how to design AI tools that preserve human agency, support authentic writing processes, and adapt to diverse user needs and disciplines (Reza et al., 2025; Hwang et al., 2025). This includes investigating the cognitive processes involved in AI-assisted writing, the role of prompt engineering, and the effects of different interaction models (e.g., co-writing, guided editing) on ownership and output quality (Picardal et al., 2025; Reza et al., 2025).

(5) Policy development and standardisation. As journal and publisher policies on AI use continue to evolve, research is

needed to analyse their consistency, effectiveness, and global adoption (Bobier et al., 2025; Veiga et al., 2025; Yoo, 2025). Future studies should also investigate the role of editorial oversight, peer review adaptation, and the development of standardised disclosure mechanisms such as machine-readable checkboxes or tiered accountability frameworks (Hosseini et al., 2025; Suchikova et al., 2025).

(6) Bias, fairness, and decolonial perspectives. Further investigation is required into how GenAI may perpetuate or amplify biases from training data, particularly regarding language, culture, and regional representation (Cambraia & Pyrrho, 2025; Cohen & Moher, 2025). A decolonial lens should be applied to examine in global knowledge production (Pretorius et al., 2025).

(7). GenAI in specialised and high-stakes domains. Research should validate the performance of GenAI and retrieval-augmented generation (RAG) systems in specialised fields such as medicine, law, and engineering (Anghelescu et al., 2025; Ghasemi et al., 2025). This includes assessing their reliability in generating accurate references, clinical guidelines, and ethical review materials, and establishing domain specific best practices for their use.

(8) Educational integration and AI literacy. There is a need to develop and evaluate AI literacy programmes for students, researchers, and educators. Future studies should explore effective pedagogical strategies for integrating GenAI into curricula while fostering critical engagement and mitigating over-reliance (Moorhouse et al., 2025; Picardal et al., 2025).

The agenda underscores the need for interdisciplinary, empirically grounded, and ethically informed research to ensure that GenAI serves as a responsible and enhancing force in academic communication.

## DISCUSSION

The discourse on AI in scholarly writing has evolved significantly from its initial reactive stance, which was predominantly cantered on the contentious issue of AI authorship following high-profile incidents involving tools like ChatGPT (Ahn, 2024; Qasem, 2023; Hufton, 2023; Lendvai, 2025). This early period of uncertainty has matured into a proactive and nuanced framework focused on responsible integration and transparency (Veiga, 2025; Yin & Chapelle, 2025), built on key pillars including transparency and disclosure, human accountability, tiered and contextual use, and collaborative Human-AI workflow (Alamri et al., 2025; Alkhawam et al., 2025; Alnaimat et al., 2025; Arar et al., 2025; Ateriya et al., 2025; Bobier et al., 2025; Daoudi, 2025; Hosseini et al., 2025; Huang et al., 2025; Hytten, 2025; Mohammadi et al., 2025; Ragel, 2025; Resnik & Hosseini, 2025; Suchikova et al., 2025; Veiga, 2025).

A critical methodological finding of this review concerns the reliability of automated bibliometric classifications in emerging, interdisciplinary fields. Our analysis revealed a significant discrepancy between the formal document types assigned by the Scopus database and the actual methodological nature of the publications, with nearly half (10 out of 18) of the review articles in our sample misclassified as primary research articles (e.g., Alamri et al., 2025; Knight, 2025; Lendvai, 2025). This carries substantial implications, as reliance on automated database filters could lead researchers to overlook a large portion of the available synthetic literature or contaminate a sample intended for primary empirical analysis, underscoring the necessity of manual, content-based verification.

The final, verified sample proved to be a key strength of this review, representing a well-balanced and methodologically diverse body of literature. The inclusion of a foundational corpus of 18 reviews (including comprehensive analyses like the hybrid systematic review by Arar et al. (2025) and the scoping review by Ghasemi et al. (2025)) provided a robust, high-level overview. This was powerfully complemented by the 37 primary articles and conceptual papers, which offered granular, empirical insights from experimental studies (Akgun et al., 2025; Erol et al., 2025), qualitative investigations (Hu et al., 2025), and critical conceptual analyses (Cambraia & Pyrrho, 2025; Hytten, 2025). This methodological triangulation ensured a comprehensive and grounded synthesis.

The thematic analysis reveals a scholarly discourse that is both complex and structurally coherent, organized around three dominant problem complexes: governance and policy, technological capabilities and risks, and systemic integrity. The strong alignment between our initial ten-theme framework and the five clusters generated by VOSviewer keyword analysis validates the core architecture of the field, confirming that governance issues like authorship and editorial policies (Yoo, 2025; Resnik & Hosseini, 2025) and technological concerns about research integrity (Akgun et al., 2025; Erol et al., 2025) form the central pillars. The algorithmic consolidation of themes further demonstrates that issues of policy, ethics, and peer review (Hosseini et al., 2025; Ebadi et al., 2025) are intrinsically intertwined in practice, reflecting a holistic institutional response.

A critical insight from the VOSviewer analysis is the field's current emphasis on control and risk mitigation over proactive educational development. The fact that educational integration (Moorhouse et al., 2025; Picardal et al., 2025) and broader socio-philosophical aspects (Cambraia & Pyrrho, 2025; Hytten, 2025) did not emerge as distinct clusters but were subsumed within themes of academic integrity and epistemic reliability, indicates a reactive phase. This suggests the community is primarily preoccupied with safeguarding existing systems of quality control and knowledge validation against AI-related disruptions, particularly through detec-

tion methodologies and integrity preservation (Erol et al., 2025; Mondal et al., 2025).

The analysis demonstrates a direct causal relationship between specific AI applications and the ethical concerns they generate. The use of AI for substantive intellectual tasks directly challenges traditional conceptions of authorship and accountability (Resnik & Hosseini, 2025; Yoo, 2025), while its application in language polishing raises distinct concerns about transparency and potential bias in scholarly evaluation, particularly for non-native English speakers (Hosseini et al., 2025; Yin & Chapelle, 2025). This application-driven nature has prompted the development of precisely targeted normative frameworks, such as the tiered disclosure model (Resnik & Hosseini, 2025) and the GAIDeT taxonomy (Suchikova et al., 2025).

A strong cross-disciplinary consensus has crystallized around several foundational principles, most notably that AI systems cannot qualify as authors and that human researchers must retain ultimate accountability for scholarly work (Yoo, 2025; Veiga, 2025). This consensus reflects a collective commitment to preserving human oversight, necessitating rigorous verification of AI outputs to prevent plagiarism and factual "hallucinations" (Anghelescu et al., 2025; Akgun et al., 2025). However, significant tensions remain in implementation, particularly regarding the balance between transparency and equity in disclosure practices. Furthermore, concerns about "algiarism" (Al Hosni, 2025), erosion of critical thinking (Raitskaya & Tikhonova, 2025), algorithmic bias (Cohen & Moher, 2025), and "technological colonialism" in access to AI tools (Cambraia & Pyrrho, 2025) highlight that the ethical challenges extend beyond immediate authorship questions to broader systemic threats.

The consolidated research agenda reveals a critical juncture in the field, marking a necessary transition from reactive problem-identification to proactive, solution-oriented inquiry (Knight, 2025; Resnik & Hosseini, 2025). The strong emphasis on developing and validating ethical frameworks and policies underscores a collective recognition that ad-hoc guidelines are insufficient. Similarly, the urgent call for more robust detection tools to address their documented unreliability (Erol et al., 2025; Mondal et al., 2025) and for longitudinal impact studies (Kumar et al., 2025; Hu et al., 2025) highlights a pressing need to gather empirical evidence on AI's tangible effects on scholarly rigor, skill development, and equity (Pretorius et al., 2025).

The agenda further highlights several underdeveloped frontiers requiring immediate scholarly attention. The focus on human-AI collaboration and user-centred design points to a significant gap in understanding the cognitive and experiential dimensions of working with AI, necessitating research that preserves human agency within collaborative workflows (Hwang et al., 2025; Picardal et al., 2025). Furthermore, the explicit calls for investigating algorithmic bias through decolonial lenses (Cambraia & Pyrrho, 2025; Cohen & Moher, 2025) and for validating AI in high-stakes domains like medicine and law (Anghelescu et al., 2025; Ghasemi et al., 2025) reveal an awareness that current systems are not context-agnostic. The inclusion of AI literacy and educational integration as a priority (Moorhouse et al., 2025; Picardal et al., 2025), while currently peripheral in the broader discourse, is crucial for ensuring that the next generation of scholars is prepared to use these tools both critically and ethically.

### Limitations

While Scopus provides extensive coverage of peer-reviewed journals and conference proceedings, several inherent limitations of the database were acknowledged. A primary constraint is its limited inclusion of grey literature, which results in the potential omission of crucial document types such as publisher and institutional policies, preprints from servers like arXiv and SSRN, and formal reports from key organizations like COPE and UNESCO, where early and influential discussions often occur. Furthermore, Scopus is known to have coverage gaps in certain regional publications and specific humanities and social science disciplines compared to other databases, which may introduce a geographic or disciplinary bias. To mitigate these limitations, supplementary strategies such as citation chasing (both forward and backward) and targeted manual searches of key organizational websites were employed.

## CONCLUSION

This 2025 scoping review has systematically mapped the landscape of literature on generative AI in scholarly writing and publishing, revealing a field characterized by rapid evolution, global engagement, and increasing methodological sophistication. Our analysis of 56 publications demonstrates a geographically diverse discourse with significant challenges in bibliometric classification, where nearly half of all review articles were misclassified in major databases that highlights the necessity of manual verification in this dynamic research area.

The discourse has matured from initial debates about AI authorship to a sophisticated understanding of how different application patterns, from substantive intellectual tasks to language enhancement, directly shape specific ethical concerns. This has led to the development of precisely targeted normative frameworks, including tiered disclosure models and AI delegation taxonomies, that acknowledge varying levels of ethical responsibility based on the nature of AI assistance.

The thematic structure of the field reveals a strong focus on governance, technological capabilities, and risk mitigation, with educational and broader cultural considerations remaining peripheral. While a strong cross-disciplinary con-

sensus has emerged around foundational principles, particularly human accountability and AI's non-author status, significant challenges persist in balancing transparency with equity, preserving scholarly rigor, and addressing systemic issues like algorithmic bias and "technological colonialism" in resource distribution.

The proposed research agenda signals a critical transition from problem-identification to solution-oriented inquiry. Priority areas include developing validated ethical frameworks, improving detection methodologies, conducting longitudinal impact studies, and advancing human-AI collaboration models. Particularly urgent is the need for research examining AI through decolonial lenses, establishing domain-specific best practices, and integrating AI literacy into educational frameworks.

As generative AI continues to transform scholarly communication, this review provides a comprehensive foundation for navigating the complex ethical landscape. The emerging frameworks and research priorities outlined here offer a pathway toward harnessing AI's potential while steadfastly upholding the core values of academic integrity, accountability, and equitable knowledge production that remain fundamental to scholarly endeavour.

## DECLARATION OF COMPETING INTEREST

None declared.

## AUTHORS' CONTRIBUTIONS

**Lilia Raitskaya**: conceptualization; data curation; formal analysis; investigation; methodology; resources; software; validation; visualization; writing – original draft; writing – re-view & editing.

**Elena Tikhonova**: conceptualization; data curation; formal analysis; investigation; methodology; resources; software; validation; visualization; writing – original draft; writing – re-view & editing.

## REFERENCES

Ahn, S. (2024). The transformative impact of large language models on medical writing and publishing: Current applications, challenges and future directions. *Korean Journal of Physiology & Pharmacology, 28*(5), 393–401. https://doi.org/10.4196/kjpp.2024.28.5.393

Akgun, M. Y., Savasci, M., Gunerbuyuk, C., Gunarar, S. O., Oktenoglu, T., Ozer, A. F., & Ates, O. (2025). Battle of the authors: Comparing neurosurgery articles written by humans and AI. *Journal of Clinical Neuroscience, 135*, Article 111152. https://doi.org/10.1016/j.jocn.2025.111152

Al Hosni, J. (2025). Preserving authorial voice in academic texts in the age of generative AI: A thematic literature review. *Arab World English Journal, 16*(3), 244–258. https://dx.doi.org/10.24093/awej/vol16no3.14

Alamri, W., Qasem, F., Alfotais, A., & Al Taisan, H. (2025). Leveraging ChatGPT AI model in academic writing and avenues for further development: SWOT framework. *Forum for Linguistic Studies*, *7*(2), 61–71. https://doi.org/10.30564/fls.v7i2.8218

Alberth, A. (2023). The use of ChatGPT in writing: A blessing or a curse in disguise? *Teflin Journal, 34*(2), 337-352. http://doi.org/10.15639/teflinjournal.v34i2/337-352

Alghazo, S., Rabab'ah, G., El-Dakhs, D. A. S., & Mustafa, A. (2025). Engagement strategies in human-written and AI-generated academic essays: A corpus-based study. *Ampersand*, *15*, Article 100237. https://doi.org/10.1016/j.amper.2025.100237

Alkhawam, M., Almobayed, A., Pandey, A., Nanda, N. C., Ebrahimi, A. J., & Ahmed, M. I. (2025). Exploring AI use policies in manuscript writing in cardiology and vascular journals. *American Heart Journal Plus: Cardiology Research and Practice*, *58*, Article 100586. https://doi.org/10.1016/j.ahjo.2025.100586

Alnaimat, F., AlSamhori, A. R. F., Hamdan, O., Seill, B., & Qumar, A. B. (2025). Perspectives of artificial intelligence use for in-house ethics checks of journal submissions. *Journal of Korean Medical Science*, *40*(21), Article e170. https://doi.org/10.3346/jkms.2025.40.e170

Anghelescu, A., Munteanu, C., Anghelescu, L. A.-M., & Onose, G. (2025). "A Midsummer Night's Dream" quest for truth: From ChatGPT "hallucinations" to RAG reasoning and ACURAI precision - A scoping review on detection, minimizing, and (almost) complete error elimination and enhancing Large Language Models' reliability. *Balneo and PRM Research Journal*, *16*(5), Article 847. https://doi.org/10.12680/balneo.2025.847

Arar, K. H., Özen, H., Polat, G., & Turan, S. (2025). Artificial intelligence, generative artificial intelligence and research integrity: A hybrid systemic review. *Smart Learning Environments*, *12*(44). https://doi.org/10.1186/s40561-025-00403-3

Archana, S. N., Renjith, V. R., Padmakumar, P. K., Shajitha, C., & Aboobaker, N. (2025). AI assisted learning and research: An exploratory study among university students and scholars. *Discover Education*, *4*(390). https://doi.org/10.1007/s44217-025-00814-x

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology, 8*, 19–32. http://doi.org/10.1080/1364557032000119616

Ateriya, N., Sonwani, N. S., Thakur, K. S., Kumar, A., & Verma, S. K. (2025). Exploring the ethical landscape of AI in academic writing. *Egyptian Journal of Forensic Sciences*, *15*(36). https://doi.org/10.1186/s41935-025-00453-1

Bobier, C., Rodger, D., & Hurst, D. (2025). Artificial intelligence policies in bioethics and health humanities: A comparative analysis of publishers and journals. *BMC Medical Ethics*, *26*(79). https://doi.org/10.1186/s12910-025-01239-9

Bozkurt, A. (2024). GenAI et al.: Cocreation, authorship, ownership, academic ethics and integrity in a time of generative AI. *Open Praxis, 16*(1), 1-10. https://doi.org/10.55982/openpraxis.16.1.654

Cambraia, L., & Pyrrho, M. (2025). Generative artificial intelligence and the risk of technological colonialism. *Frontiers in Political Science*, *7*, Article 1628139. https://doi.org/10.3389/fpos.2025.1628139

Cheng, A., Calhoun, A., & Reedy, G. (2025). Artificial intelligence-assisted academic writing: Recommendations for ethical use. *Advances in Simulation, 10*(22). https://doi.org/10.1186/s41077-025-00350-6

Cohen, J. F., & Moher, D. (2025). Generative artificial intelligence and academic writing: Friend or foe? *Journal of Clinical Epidemiology, 179*, Article 111646. https://doi.org/10.1016/j.jclinepi.2024.111646t6

Comas-Forgas, R., Koulouris, A., & Kouis, D. (2025). 'AI-navigating' or 'AI-sinking'? An analysis of verbs in research articles titles suspicious of containing AI-generated/assisted content. *Learned Publishing*, *38,* Article e1647. https://doi.org/10.1002/ieap.1647

Daoudi, M. (2025). Ethical limits and suggestions for improving the use of AI in scientific research, academic publishing, and the peer review process, based on deontological and consequentialist viewpoints. *Discover Education*, *4*, 241. https://doi.org/10.1007/s44217-025-00696-z

Ebadi, S., Nejadghanbar, H., Salman, A. R., & Khosravi, H. (2025). Exploring the impact of generative AI on peer review: Insights from journal reviewers. *Journal of Academic Ethics*, *23*, 1383–1397. https://doi.org/10.1007/s10805-025-09604-4

Erdat, E. C., & Çay Şenler, F. (2025). Turkish medical oncologists' perspectives on integrating artificial intelligence: Knowledge, attitudes, and ethical considerations. *BMC Medical Ethics*, *26*(1), Article 82. https://doi.org/10.1186/s12910-025-01249-7

Erel, S., Erkocak Arabaci, O., & Pampal, H. K. (2025). Examining the frequency of artificial intelligence generated content in anesthesiology and intensive care journal publications: A cross sectional study. *Medicine, 104*(8), Article e41594. https://doi.org/10.1097/MD.0000000000041594

Erol, G., Ergen, A., Erol, B. G., Ergen, Ş. K., Bora, T. S., Çölgeçen, A. D., Araz, B., Şahin, C., Bostancı, G., Kilç, İ., Macit, Z. B., Sevgi, U.T., & Güngor, A. (2025). Can we trust academic AI detective? Accuracy and limitations of AI-output detectors. *Acta Neurochirurgica*, *167*(1), Article 214. https://doi.org/10.1007/s00701-025-06622-4

Fabiano, N., Gupta, A., Bhambra, N., Luu, B., Wong, S., Maaz, M., Fiedorowicz, J. G., Smith, A. L., & Solmi, M. (2024). How to optimize the systematic review process using AI tools. *JCPP Advances*, *4*(2), Article e12234. https://doi.org/10.1002/jcv2.12234

Fakharifar, A., Beizavi, Z., Pouramini, A., & Haseli, S. (2025). Application of artificial intelligence and ChatGPT in medical writing: A narrative review. *Journal of Medical Artificial Intelligence*, *8*, Article 52. https://doi.org/10.21037/jmai-24-342

Fung, C.-H., & Ng, S.-P. (2025). AI-driven simulation-based research: A new frontier in research innovation. *Computers and Education: Artificial Intelligence, 9*, 100453. https://doi.org/10.1016/j.caeai.2025.100453

Ghasemi, S. F., Amiri, P., & Galavi, Z. (2025). Advantages and limitations of ChatGPT in healthcare: A scoping review. *Health Science Reports*, *8*, Article e71219. https://doi.org/10.1002/hsr2.71219

Goyibova, N., Muslimov, N., Kannazarova, Z., Kadirova, N., Alautdinova, K., & Ismatullaeva, I. (2025). Exploring the impact of artificial intelligence on academic writing: A bibliometric analysis of trends, advancements, and ethical challenges. *Forum for Linguistic Studies*, *7*(6), 342–360. https://doi.org/10.30564/fls.v7i6.9054

Hosseini, M., Gordijn, B., Kaebnick, G. E., & Holmes, K. (2025). Disclosing generative AI use for writing assistance should be voluntary. *Research Ethics*, *21*(4), 728-735. https://doi.org/10.1177/17470161241212345

Hu, H., Zhou, Q., & Hashim, H. (2025). Negotiating identity in the age of ChatGPT: Non-native English researchers' experiences with AI-assisted academic writing. *Humanities and Social Sciences Communications*, *12*(1), Article 965. https://doi.org/10.1057/s41599-025-05351-4

Huang, W., Liang, Y., Wei, X., & Du, Y. (2025). Ophthalmology journals' guidelines on generative artificial intelligence: A comprehensive analysis. *American Journal of Ophthalmology, 271*, 445–454. https://doi.org/10.1016/j.ajo.2024.12.023

Hufton, A.L. (2023). No artificial intelligence authors, for now. *Patterns, 4*, 14, 2023. https://doi.org/10.1016/j.patter.2023.100731

Hwang, A. H.-C., Liao, Q. V., Blodgett, S. L., Olteanu, A., & Trischler, A. (2025). «It was 80% me, 20% AI»: Seeking authenticity in co-writing with large language models. *Proceedings of the ACM on Human-Computer Interaction*, 9(2), Article CSCW122, 1–41. https://doi.org/10.1145/3711020

Hytten, K. (2025). Ethics, AI, and irresistible temptations. *Educational Theory*, *75*(1), 1–8. https://doi.org/10.1111/edth.12635

Khalifa, M., & Albadawy, M. (2024). Using artificial intelligence in academic writing and research: An essential productivity tool. *Computer Methods and Programs in Biomedicine Update, 5*, Article 100145. https://doi.org/10.1016/j.cmpbup.2024.100145

King, M. R., & ChatGPT. (2023). A conversation on artificial intelligence, chatbots, and plagiarism in higher education. *Cellular and Molecular Bioengineering, 16*, 1–2. https://doi.org/10.1007/s12195-022-00754-8

Knight, S. (2025). Understanding use of evidence in AI ethics guidelines development through a PRISMA-ETHICS informed scoping review of guidelines. *Computers and Education Open, 9*, Article 100281. https://doi.org/10.1016/j.caeo.2025.100281

Knöchel, T.-D., Schweizer, K. J., Acar, O. A., Akü, A. M., Al-Hoorie, A. H., Buehler, F., Elsherif, M. M., Giannini, A., Heyselaar, E., Hosseini, M., Ilangovan, V., Kovacs, M., Lin, Z., Liu, M., Peeters, A., van Ravenzwaaij, D., Vranka, M. A., Yamada, Y., Yang, Y.-F., & Aczel, B. (2025). Core principles of responsible generative AI usage in research. *AI and Ethics*, *5*, 6371–6377. https://doi.org/10.1007/s43681-025-00768-8

Kumar, R., Arun Kumar, P., Anbu Selvam, A., Dhamu, I., & Balasubramanian, N. (2025). The impact of ChatGPT on orthopaedic residents' scientific writing: A quasi-experimental study. *Journal of Clinical Orthopaedics and Trauma, 69*, Article 103116. https://doi.org/10.1016/j.jcot.2025.103116

Lendvai, G. F. (2025). ChatGPT in academic writing: A scientometric analysis of literature published between 2022 and 2023. *Journal of Empirical Research on Human Research Ethics*, *20*(3), 131–148. https://doi.org/10.1177/15562646251350203

Lim, B., Seth, I., Cevik, J., Mu, X., Sofiadellis, F., Cuomo, R., & Rozen, W. M. (2025). Artificial intelligence tools in surgical research: A narrative review of current applications and ethical challenges. *Surgeries, 6*(3), 55. https://doi.org/10.3390/surgeries6030055

Lim, G. H., Tan, M. L., Hoe, V. C. W., & Koh, D. (2025). Generative AI in peer review process for occupational health. *Occupational Medicine, 75*(6), 242–249. https://doi.org/10.1093/ocmed/kqa051

Liu, Y., Kong, W., & Kaygac, M. (2025). ChatGPT applications in academic writing: a review of potential, limitations, and ethical challenges. *Arquivos Brasileiros de Oftalmologia*, *88*(3), Article e2024-0269. https://doi.org/10.5935/0004-2749.2024-0269

Miller, A. S., Tyagi, A., Sudah, S. Y., Rompala, A., Nicholson, A. D., Srikumaran, U., & Menendez, M. E. (2025). Evaluation of the impact of large language learning models on publications in the Journal of Shoulder and Elbow Surgery. *JSES International, 9*, 1803–1808. https://doi.org/10.1016/j.jseint.2024.10.001

Mo, Z., & Crosthwaite, P. (2025). Exploring the affordances of generative AI large language models for stance and engagement in academic writing. *Journal of English for Academic Purposes, 75*, Article 101499. https://doi.org/10.1016/j.jeap.2025.101499

Mohammadi, E., Thelwall, M., Cai, Y., Collier, T., Tahamtan, I., & Eftekhar, A. (2025). Is generative AI reshaping academic practices worldwide? A survey of adoption, benefits, and concerns. *Information Processing and Management*, *63*, Article 104350. https://doi.org/10.1016/j.ipm.2025.104350

Mohan, C., Singhal, M., & Verma, A. (2025). Navigating the editorial gauntlet for excellence in radiology publishing. *Indian Journal of Radiology and Imaging*, *35*(Suppl. S1), S102–S109. https://doi.org/10.1055/s-0044-1801269

Mondal, H., Mondal, S., & Jana, S. (2025). The artificial intelligence dilemma in academic writing: Balancing efficiency and integrity. *Indian Journal of Cardiovascular Disease in Women*, *10*(3), 225–230. https://doi.org/10.25259/IJCDW_86_2024

Moorhouse, B. L., Nejadghanbar, H., & Yeo, M. A. (2025). Study quality in the age of AI: A disciplinary framework for using GenAI in TESOL research. *TESOL Quarterly*, *0*(0), 1–15 (in press). https://doi.org/10.1002/tesq.70026

Nassar, J. E., Farias, M. J., Singh, M., Dinh, P. V., Sahhar, M., Daher, M., Fallon, R., Marcaccio, S. E., Daniels, A. H., & Owens, B.D. (2025). Large language model–based writing in published sports medicine research: Uncovering a growing influence. *The Orthopaedic Journal of Sports Medicine*, *13*(9), Article 23259671251371234. https://doi.org/10.1177/23259671251371234

Öztürk, A., Karahan, A. T., Günay, S., Erdal, A. S., Komut, S., Komut, E., & Yiğit, Y. (2025). Artificial intelligence as author: Can scientific reviewers recognize GPT-4o-generated manuscripts? *American Journal of Emergency Medicine*, *97*, 216–219. https://doi.org/10.1016/j.ajem.2025.07.034

Parlak, L., Levendoglu, F., Balevi Batur, E., Akyildiz Tezcan, E., & Albayrak Gezer, I. (2025). Artificial or authentic? A comparative study of article titles generated by humans and AI in physical medicine and rehabilitation. *Genel Tip Dergisi, 35*(4), 775–781. https://doi.org/10.54005/geneltip.1725576

Picardal, M. T. (2025). Utilization of AI-driven smart prompts in academic research in higher education institutions. *International Journal of Learning, Teaching and Educational Research*, *24*(9), 384–404. https://doi.org/10.26803/ijlter.24.9.19

Prakash, A., Aggarwal, S., Varghese, J. J., & Varghese, J. J. (2025). Writing without borders: AI and cross-cultural convergence in academic writing quality. *Humanities & Social Sciences Communications*, *12*(1058). https://doi.org/10.1057/s41599-025-05484-6

Premat, C., & Farazouli, A. (2025). Academic integrity vs. artificial intelligence: A tale of two AIs. *Práxis Educativa, 20,* Article e24871. https://doi.org/10.18226/e24871.v20

Pretorius, L., Huynh, H.-H., Pudyanti, A. A. A. R., Li, Z., Noori, A. Q., & Zhou, Z. (2025). Empowering international PhD students: Generative AI, Ubuntu, and the decolonisation of academic communication. *The Internet and Higher Education, 67*, Article 101038. https://doi.org/10.1016/j.iheduc.2025.101038

Qasem, F. (2023). ChatGPT in scientific and academic research: Future fears and reassurances. *Library Hi Tech News, 40*(3), 30-32. https://doi.org/10.1108/LHTN-03-2023-0043

Radtke, A., & Rummel, N. (2025). Generative AI in academic writing: Does information on authorship impact learners' revision behavior? *Computers and Education: Artificial Intelligence, 8*, Article 100350. https://doi.org/10.1016/j.caeai.2024.100350

Rafi, M. S., & Amjad, I. (2025). The role of generative AI in writing doctoral dissertation: Perceived opportunities, challenges, and facilitating strategies to promote human agency. *Discover Education, 4*(165). https://doi.org/10.1007/s44217-025-00503-9

Ragel, R. (2025). Embracing AI in scholarly publishing: Enhancing integrity and expanding access. *Journal of the National Science Foundation of Sri Lanka, 53*(1), 1–2. http://dx.doi.org/10.4038/jnsfsr.v53i1.12647

Raitskaya, L., & Tikhonova, E. (2024). Appliances of generative AI-powered language tools in academic writing: A scoping review. *Journal of Language and Education*, *10*(4), 5–30. https://doi.org/10.17323/jle.2024.24181

Raitskaya, L., & Tikhonova, E. (2025). Enhancing critical thinking skills in ChatGPT-human interaction: A scoping review. *Journal of Language and Education, 11*(2), 5–19. https://doi.org/10.17323/jle.2025.27387

Resnik, D. B., & Hosseini, M. (2025). Disclosing artificial intelligence use in scientific research and publication: When should disclosure be mandatory, optional, or unnecessary? *Accountability in Research*. Advance online publication. https://doi.org/10.1080/08989621.2025.2481949

Reza, M., Thomas-Mitchell, J., Dushniku, P., Laundry, N., Williams, J. J., & Kuzminykh, A. (2025). Co-writing with AI, on human terms: Aligning research with user demands across the writing process. *Proceedings of the ACM on Human-Computer Interaction*, *9*(7), Article CSCW385. https://doi.org/10.1145/3757566

Sparkman, M., & Witt, A. (2025). Claude AI and literature reviews: An experiment in utility and ethical use. *Library Trends*, *73*(3), 355–380. https://doi.org/10.1353/lib.2025.a961199

Suchikova, Y., Tsybuliak, N., Teixeira da Silva, J. A., & Nazarovets, S. (2025). GAIDeT (Generative AI Delegation Taxonomy): A taxonomy for humans to delegate tasks to generative artificial intelligence in scientific research and publishing. *Accountability in Research*. Advance online publication. https://doi.org/10.1080/08989621.2025.2544331

Tikhonova E., & Raitskaya L. (2023). ChatGPT: Where Is a Silver Lining? Exploring the realm of GPT and large language models. *Journal of Language and Education*, *9*(3), 5-11. https://doi.org/10.17323/jle.2023.18119

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., Straus, S.E. (2018). PRISMA extension for scoping reviews (PRISMA- ScR): Checklist and explanation. *Annals of Internal Medicine , 169*(7), 467–473. https://doi.org/10.7326/M18-0850

Veiga, A. (2025). Ethical guidelines for the use of generative artificial intelligence and artificial intelligence-assisted tools in scholarly publishing: A thematic analysis. *Science Editing, 12*(1), 28–34. https://doi.org/10.6087/kcse.2025.12.1.28

Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., & Gašević, D. (2024). Practical and ethical challenges of large language models in education: A systematic scoping review. *British Journal of Educational Technology*, *55*(1), 90–112. https://doi.org/10.1111/bjet.13370

Yaqin, L. N., Hassan, H., & Yusof, B. (2025). Performance and accuracy of ChatGPT in generating Malay academic texts: A comparative study with expert corrections. *LLT Journal: A Journal on Language and Language Teaching*, *28*(1), 495–517. https://doi.org/10.24071/llt.v28i1.11698

Yin, S., & Chapelle, C. A. (2025). A systematic examination of generative artificial intelligence (GenAI) use guidelines in applied linguistics journals. *Research Methods in Applied Linguistics, 4*, 100227. https://doi.org/10.1016/j.rmal.2025.100227

Yoo, J.-H. (2025). Defining the boundaries of AI use in scientific writing: A comparative review of editorial policies. *Journal of Korean Medical Science, 40*(23), Article e187. https://doi.org/10.3346/jkms.2025.40.e187

# APPENDIX 1

## Reviews Included in the Review

| Nos | Reference | Type of Review, Details |
|---|---|---|
| 1. | Arar et al., 2025 | Hybrid Systemic Review combining:<br>Bibliometric Analysis and Thematic Content Analysis; 595 documents |
| 2. | Ateriya et al., 2025 | Narrative Literature Review; the number of publications are not explicitly stated |
| 3. | Fakharifar et al., 2025 | Narrative Review; the number of publications is not explicitly stated |
| 4. | Ghasemi et al., 2025 | Scoping review; 28 studies |
| 5. | Goyibova et al., 2025 | Bibliometric Analysis; 108 documents |
| 6. | Mohan et al., 2025 | Not applicable; no methodological framework |
| 7. | Mondal et al., 2025 | A perspective review; the number of publications is not explicitly stated |
| 8. | Raitskaya & Tikhonova, 2025 | A scoping review; 30 empirical studies |

# APPENDIX 2

## Articles Included in the Review

| Nos | Reference | Research Design | Population |
|---|---|---|---|
| 1 | Akgun et al., 2025 | Randomized, blinded non-inferiority trials | 20 neurosurgeons assessing 10 neurosurgery articles (5 human-written, 5 AI-generated) |
| 2 | Al Hosni, 2025 | Not applicable<br>Though marked as “Article” in the Scopus, it is a thematic literature review | 18 scholarly papers (published between 2023 and 2025) |
| 3 | Alamri et al., 2025 | Not applicable<br>Though marked as “Article” in the Scopus, it is a Systematic Review combined with a SWOT Analysis Framework | 23 research papers (published in 2024) |
| 4 | Alghazo et al., 2025 | A corpus-based, mixed-methods approach | 80 essays |
| 5 | Alkhawam et al., 2025 | A cross-sectional study | Not applicable |
| 6 | Alnaimat et al., 2025 | A narrative/perspective review | 213 MEDLINE-indexed cardiovascular journals |
| 7 | Anghelescu et al., 2025 | Though marked as “Article” in the Scopus, it is a scoping review | 78 papers |
| 8 | Archana et al., 2025 | A cross-sectional online survey | 421 post-graduate (PG) students and research scholars (PhD) from Science, Social Science, and Technology disciplines |
| 9 | Bobier et al., 2025 | A comparative policy analysis | 50 bioethics and health/medical humanities journals |
| 10 | Cambraia & Pyrrho, 2025 | A conceptual analysis | Not applicable |
| 11 | Cohen & Moher, 2025 | A commentary/article, not an empirical study | Not applicable |
| 12 | Comas-Forgas et al., 2025 | A mixed-methods study integrating qualitative observation with a bibliometric analysis | Research article titles indexed in the Scopus database from 2015 to 2024 |
| 13 | Daoudi, 2025 | A mixed-methods approach using a structured questionnaire with quantitative (closed-ended) and qualitative (open-ended) components | 42 researchers from various scientific disciplines |
| 14 | Ebadi et al., 2025 | A qualitative research design using an open-ended questionnaire | 12 journal reviewers from diverse academic disciplines |
| 15 | Erel et al., 2025 | An observational cross-sectional study | 1268 articles (abstracts) from 86 journals in the “Anesthesiology” and “Anesthesiology and Intensive Care” categories indexed in Web of Science |
| 16 | Erol et al., 2025 | An experimental study | 1,000 text samples (250 human-authored abstracts/introductions from pre-ChatGPT era neurosurgery articles and 750 AI-generated counterparts) |
| 17 | Fung & Ng, 2025 | A position paper with a conceptual case study | Not applicable |
| 18 | Hosseini et al., 2025 | An ethical argument and policy analysis paper | Not applicable |
| 19 | Hu et al., 2025 | A qualitative exploratory design | 25 non-native English-speaking researchers from various countries and academic disciplines |
| 20 | Huang et al., 2025 | A cross-sectional bibliometric analysis | 140 ophthalmology journals listed in the Scimago Journal and Country Rankings |
| 21 | Hwang et al., 2025 | A mixed-methods study | Part 1: 19 professional writers. Part 2: 30 avid readers |

| Nos | Reference | Research Design | Population |
|---|---|---|---|
| 21 | Hytten, 2025 | A philosophical reflective essay | Not applicable |
| 22 | Knight, 2025 | A scoping review | 592 distinct items were identified, with 47 guideline-sets grouped as the final units of analysis (75 documents in total) |
| 23 | Kumar et al., 2025 | A prospective, quasi-experimental, single-cohort study | 36 orthopedic junior residents from a tertiary care medical college and hospital |
| 24 | Lendvai, 2025 | A scientometric review using Document Co-citation Analysis (DCA) | 171 peer-reviewed Scopus-indexed articles |
| 25 | Lim G.H. et al., 2025 | A two-phase experimental study | Eight occupational health manuscripts (five research articles, one review, one commentary, one case report); reviewers included three AI tools and five human Occupational Medicine residents |
| 26 | Liu et al., 2025 | A systematic scoping review marked in the Scopus database as “Article” | 327 documents |
| 27 | Miller et al., 2025 | A comparative analysis of published articles | 232 publications from the Journal of Shoulder and Elbow Surgery |
| 28 | Mo & Crosthweite, 2025 | A comparative corpus analysis of human-written and AI-generated academic essays | 30 human-written essays from the British Academic Written English (BAWE) corpus and 90 AI-generated essays (30 from each of the three LLMs) |
| 29 | Moorhouse et al., 2025 | Not applicable | Not applicable |
| 30 | Nassar et al., 2025 | Cross-sectional study | All research articles published from January 2023 to December 2024 in the five sports medicine journals |
| 31 | Öztürk et al., 2025 | Experimental study / Diagnostic accuracy study | 14 reviewers |
| 32 | Parlak et al., 2025 | Comparative survey study | 15 academic experts in Physical Medicine and Rehabilitation with an H-index ≥3 |
| 33 | Picardal et al., 2025 | Qualitative descriptive design | 65 preservice science teachers (BSED Science) from a teacher training institution in Central Visayas, Philippines |
| 34 | Prakash et al., 2025 | Longitudinal large-scale quantitative analysis | Academic abstracts from research articles listed under the Social Sciences Citation Index (SSCI) published by three major academic publishers between 2012 and 2024 |
| 35 | Premat et al., 2025 | A qualitative phenomenographic analysis | 42 students enrolled in a course on academic integrity |
| 36 | Pretorius et al., 2025 | Collaborative Autoethnography | Five international PhD students and the authors |
| 37 | Radtke & Rummel, 2025 | An experimental study with a mixed between- and within-subjects design | 303 participants |
| 38 | Rafi & Amjad, 2025 | Qualitative; survey and test with open-ended questions | 86 doctoral scholars and 7 thesis supervisors from social and human sciences disciplines |
| 39 | Resnik & Hosseini, 2025 | A conceptual/commentary paper | Not applicable |
| 40 | Reza et al., 2025 | Mixed methods; A systematic literature review | A final dataset of 109 HCI papers involving 15 writers |
| 41 | Sparkman & Witt, 2025 | A comparative experiment | A set of scholarly articles |
| 42 | Suchikova et al., 2025 | A conceptual paper; a structured exploratory literature review | Not applicable |
| 43 | Veiga et al., 2025 | Not applicable | Not applicable |
| 44 | Yaqin et al., 2025 | A comparative research design | 40 passages from peer-reviewed Malay journal articles |
| 45 | Yin & Chapelle, 2025 | Text analysis / documentary analysis, a policy review | 76 journal policy documents (from the 170 journals analyzed) |
| 46 | Yoo, 2025 | A comparative policy review | Not applicable |

# APPENDIX 3

## Key Findings and Research Agendas of Publications included in the Review

| Reference | Type of GenAI / AI-Powered Tools | Key Findings | Further Research Agenda |
|---|---|---|---|
| Akgun et al., 2025 | Generative AI: ChatGPT (specifically the GPT-4 model) | No significant differences in overall quality, impression, criteria satisfaction, or degree of detail between AI and human articles.<br>AI-generated articles had significantly higher readability scores (Lix score and Flesch-Kincaid grade level).<br>Expert assessors could only correctly identify authorship 61% of the time.<br>Preferences were almost evenly split (47% preferred AI, 44% preferred human, 9% no preference).<br><br>AI is capable of producing comparable, highly readable scientific content that is often indistinguishable from human-written work | Refine ChatGPT's capabilities for medical writing.<br>Develop stronger AI-detection algorithms for journal submissions.<br>Establish and update ethical guidelines for AI use in research writing.<br>Explore the balance between AI innovation and accountability to maintain literature credibility. |
| Al Hosni, 2025 | Generative AI: ChatGPT (versions 3.5 & 4), Copilot, Gemini, Claude | AI-generated texts are grammatically polished, structurally coherent, and have higher lexical diversity and syntactic complexity.<br>However, they lack emotional depth, rhetorical nuance, personal engagement, and critical insight.<br>AI texts underutilize hedges, boosters, and engagement markers, resulting in a detached, neutral, and formulaic tone.<br>Human-authored texts show greater specificity, cultural context, authentic emotional depth, and signs of critical thinking.<br>Over-reliance on AI can lead to the erosion of a student's authorial voice and linguistic fingerprint, a phenomenon linked to "Algiarism." | Examine the long-term impact of AI tools on voice development.<br>Explore voice expression across different genres and disciplines.<br>Investigate student perceptions of authorship and identity in AI-assisted writing |
| Alamri et al., 2025 | Primary Tool: ChatGPT (various versions, including GPT-3.5, GPT-4).<br>Other Tools: Google Bard/ Gemini, Google PaLM 2, Anthropic's Claude 2, Quillbot, Grammarly, Bing AI/Chat, DeepL | Strengths:<br>• Generates ideas and assists with writing planning.<br>• Improves efficiency and writing quality.<br>• Provides personalization and immediate feedback/support.<br>Weaknesses:<br>• Leads to user overreliance.<br>• Raises ethical concerns (e.g., plagiarism).<br>• Can provide incorrect information and sources.<br>Opportunities:<br>• Enables enhanced learning strategies.<br>• Facilitates collaborative learning.<br>• Drives innovative pedagogical practices.<br>Threats:<br>• Causes degradation of critical and creative thinking skills.<br>• Creates issues due to lack of access to technology/the internet.<br>• Raises data privacy concerns | Not explicitly stated. The conclusion implies a need for future research on developing safe and ethical methods for using AI and on the practical implementation of AI technologies in educational settings |

| Reference | Type of GenAI / AI-Powered Tools | Key Findings | Further Research Agenda |
|---|---|---|---|
| Alghazo et al., 2025 | ChatGPT (Generative AI language model) | Human writers use significantly more Expansion and Contraction strategies than AI.<br>Humans use more Entertain and Disclaim markers, showing nuanced argumentation and active opposition to counterarguments.<br>AI relies heavily on hedging and neutral citation (Acknowledge), with minimal use of Proclaim markers and no use of Concur.<br>AI texts are more expository and less persuasive, avoiding strong claims or critical engagement with sources.<br>Statistical tests (chi-square) confirm significant differences in engagement strategy use between human and AI texts | Explore how AI can be designed to adapt engagement strategies to different academic disciplines or genres.<br>Investigate the integration of emotional and ethical dimensions into AI-generated texts.<br>Conduct longitudinal studies to track improvements in AI writing capabilities over time. |
| Alkhawam et al., 2025 | Large Language Models (LLMs), Generative AI, ChatGPT (specifically mentioned) | 79.8% of the 213 cardiovascular journals had an AI policy.<br>A higher proportion of vascular journals (89%) had AI policies compared to cardiology journals (78%).<br>98.7% of journals with an AI policy explicitly excluded AI from being listed as an author.<br>100% of journals with an AI policy allowed the use of AI-assisted content.<br>The presence of an AI policy was not significantly associated with Journal Impact Factor (JIF) or other journal metrics (CiteScore, SNIP, SJR, JCI). | Track changes in AI policies over time due to the rapid evolution of the technology.<br>Broaden research to include non-English journals and other scientific disciplines beyond cardiology and vascular medicine.<br>Explore other influencing variables, such as publisher practices or editorial board decisions, on AI policy adoption. |
| Alnaimat et al., 2025 | Bias Detection: YesChat.ai<br>Plagiarism & AI-Generated Content Detection: Originality.AI, GPTZero, Scribbr AI Detector, iThenticate (Turnitin), Viper, Grammarly, Duplichecker<br>Image Integrity: Proofig, ImageRights, TinEye<br>Journal Submission & Review Platforms: Elsevier's Editorial Manager, AIRA (Frontiers), Enago Read<br>Citation Analysis: CIDRE algorithm, Scite's Reference Check<br>Authorship Analysis (Stylometry): SciDetect | AI can enhance the efficiency, consistency, and integrity of in-house ethics checks for journal submissions.<br>Specific AI tools are available for detecting plagiarism, AI-generated text, image manipulation, bias, missing ethics statements, citation manipulation, and authorship issues.<br>AI can be integrated into submission systems (e.g., Editorial Manager, AIRA) to automate compliance checks.<br>AI tools for detecting conflicts of interest and performing stylometric analysis for authorship are emerging but require further development.<br>Human oversight remains crucial to interpret AI findings and make final ethical judgments | Enhance AI tools for detecting authorship issues and conflicts of interest.<br>• Improve the accuracy and reduce false-positive/false-negative rates of AI detectors.<br>• Explore the integration of AI tools directly into manuscript submission workflows.<br>• Investigate the ethical implications and potential misuse of AI in peer review (e.g., confidentiality breaches, fake citations). |

| Reference | Type of GenAI / AI-Powered Tools | Key Findings | Further Research Agenda |
|---|---|---|---|
| Anghelescu et al., 2025 | Base LLMs: GPT-3.5, GPT-4, GPT-4 Turbo, Anthropic Claude 2, Google Bard/ Gemini, Mixtral, Command R+<br>Enhancement Frameworks: Retrieval-Augmented Generation (RAG), ACURAI<br>Specific Systems: NotebookLM, iDISK 2.0, BiomedRAG, MKRAG, RadioRAG | LLMs alone are prone to hallucinations and provide unreliable information in specialized domains like medicine.<br>RAG-enhanced models significantly improve accuracy and reliability by grounding responses in real-time, verified external sources (e.g., scientific databases).<br>RAG improved performance odds by an overall ratio of 1.35 and increased accuracy in medical question-answering from 44.46% to 48.54%.<br>The ACURAI framework, based on "phrase dominance and discrete functional units (DFUs)," demonstrated a 91-100% elimination of junk outputs in GPT-3.5 and GPT-4 when tested on the "RAG-Truth Dataset."<br>RAG-LLMs show high application potential in medical education, clinical decision support, radiology, healthcare administration, and patient/caregiver education | Broader systematic reviews or empirical testing of ACURAI in clinical settings.<br>Future advancements in LLM performance using RAG techniques, fine-tuning, and reinforcement learning.<br>Addressing the challenges of implementation, scalability, latency, and computational consumption in ACURAI-RAG systems |
| Arar et al., 2025 | GenAI: ChatGPT, Bard, Claude, GPT models.<br>AI Tools: AI-based diagnostic systems (e.g., CAD4TB, qXR, Lunit), drug discovery models, educational personalization tools, AI auditors, creative art algorithms, blockchain-integrated AI. | Rapid increase in publications on GenAI & research integrity post-2020.<br>China, USA, and India are the most productive countries.<br>Key research foci include ethical guidelines and transparency.<br>Major opportunities are enhanced efficiency and personalized learning.<br>Major threats are algorithmic bias and academic dishonesty.<br>Research integrity requires technical standards and ethical guidelines | Develop ethical principles integrated with oversight systems.<br>Form international AI ethics policies and interdisciplinary cooperation.<br>Investigate the balance between AI benefits and socio-ethical risks.<br>Create extensive regulations for AI/GenAI implementation across sectors |
| Archana et al., 2025 | GenAI: ChatGPT, Google Bard/Gemini, Microsoft Bing/Copilot AI, Ask AI, Elicit, Consensus, Scispace, ChatPDF, Gamma, Slides AI, GitHub Copilot, OpenAI Codex.<br>Traditional AI: Grammarly, Quillbot, Trinka, Grammar Check, Scribbr, Research Rabbit, Litmaps, Scholarcy | Over 90% of respondents use AI tools for academic activities. ChatGPT is the most popular tool, used for four different activities. The most common uses are for understanding/learning concepts, grammar checking, and academic writing. A significant difference in AI adoption exists across academic levels (PG vs. PhD). Discipline influenced AI usage for specific tasks like plagiarism detection and literature reviewing. No significant difference in AI tool usage was found across genders. 9.7% of non-users cited unawareness, ethical concerns, and trust issues as reasons | Incorporate multivariate regression models to refine insights. Explore discipline-specific barriers to AI adoption. Conduct focused surveys or in-depth interviews across disciplines. Examine the long-term academic impact of AI on learning outcomes and critical thinking skills. Investigate the effectiveness of AI training programs in universities |
| Ateriya et al., 2025 | Large Language Models (LLMs): ChatGPT, GPT-3, GPT-4, Gemini.<br>Other AI Tools: AI-powered plagiarism detection software, AI systems for peer review and reviewer selection | AI integration in academic writing raises major ethical concerns regarding authorship, intellectual property, and transparency. AI can perpetuate biases from its training data, leading to unfair outcomes. The use of AI increases the risk of both intentional and unintentional plagiarism. Data privacy is a critical concern when AI handles sensitive information. Human oversight remains essential to ensure accountability and integrity in AI-assisted work. AI can enhance the peer review process but requires careful management to avoid bias | Develop comprehensive ethical guidelines and frameworks for AI use in academia.<br>Create and implement effective strategies to identify and mitigate bias in AI algorithms. Investigate and establish protocols for AI transparency and accountability. Design education and training programs for researchers on the ethical use of AI. Study and promote equitable access to AI tools across institutions to bridge resource gaps. |

| Reference | Type of GenAI / AI-Powered Tools | Key Findings | Further Research Agenda |
|---|---|---|---|
| Bobier et al., 2025 | Generative AI, large language models (LLMs), ChatGPT, AI-assisted writing tools | A minority of journals had a clear AI policy, while over half were published by a publisher with a policy. Publisher policies were generally favourable to AI-assisted manuscripts, but a small number of journals explicitly prohibited AI-generated text. There was significant inconsistency in disclosure requirements and the extent of permissible AI use. The lack of standardized guidelines creates confusion for authors and risks to scholarly integrity | Exploring editorial adherence to publisher-level policies, investigating editors' perspectives on AI-assisted manuscripts, and developing consensus guidelines through professional societies like the American Society of Bioethics and Humanities |
| Cambraia & Pyrrho, 2025 | Generative AI, Large Language Models (LLMs), ChatGPT | GenAI poses a risk of technological colonialism by re-enacting colonial power dynamics. It centralizes knowledge production and enunciation, perpetuates global inequities, and exploits data and labour from the Global South. High-level ethical principles are insufficient to address these challenges, requiring a decolonial framework focused on the power dynamics of knowledge enunciation | Examining how discourse is artificially generated and why humans are convinced by it; Investigating the role of epistemic dynamics in digital colonialism; Exploring how diverse cosmovisions can influence technological design |
| Cohen & Moher, 2025 | Generative AI, Large Language Models (LLMs), ChatGPT, GPT-3.5, Meta's Llama 2 | Generative AI can assist in overcoming language barriers, expediting the writing process, and potentially improving adherence to reporting guidelines. Significant threats include the perpetuation of biases from training data, the generation of inaccurate information and fabricated references ("hallucinations"), complex authorship and accountability issues, potential exacerbation of inequalities among researchers, and the risk of fuelling paper mills and low-value publications | Research into whether GenAI is prone to distorting research findings ("spin"); Applications of GenAI to better implement reporting guidelines; Development of harmonized guidance for reporting GenAI use. |
| Comas-Forgas et al., 2025 | Generative AI (GAI), Large Language Models (LLMs), ChatGPT (specifically GPT-4.0) | A marked increase in the usage of 15 specific verbs (e.g., revolutionizing, unveiling, enhancing) was observed in academic titles from 2023-2024, suggesting AI's influence. The mean Year-over-Year Growth (YoYG) for these verbs was 99.9%. The trend was most pronounced in multidisciplinary studies, health sciences, and sciences/engineering, and among authors from non-English speaking countries. An estimated 63,780 titles over 2023-2024 were potentially AI-generated or assisted | Expanding the analysis to include more linguistic markers beyond the 15 verbs; Incorporating data from additional databases beyond SCOPUS; Examining individual articles for explicit author disclosures on GAI tool usage and how such disclosures affect academic integrity |

| Reference | Type of GenAI / AI-Powered Tools | Key Findings | Further Research Agenda |
|---|---|---|---|
| Daoudi, 2025 | ChatGPT, Copilot, Napkin, Meta AI, Perplexity, Bing, Gemini, Leonardo AI | AI is viewed as a powerful tool for enhancing research efficiency, particularly in data analysis (61.9%) and hypothesis generation (45.2%). Significant ethical concerns include bias, transparency, and trust. In academic writing, 81% support AI use with proper acknowledgment, and 76.2% are open to AI-assisted peer review under human supervision. The future role of AI is seen as complementary to human expertise (69%). Key barriers include limited access to AI tools (47.6%), high costs (38.1%), and insufficient technical skills (45.2%) | Broader sampling strategies involving larger and more demographically, institutionally, and geographically diverse populations; Triangulating self-reported data with observational studies or document analyses |
| Ebadi et al., 2025 | Large Language Models (LLMs), specifically mentioned as "generative AI tools" and "AI-based chatbots" (e.g., tools like ChatGPT are implied) | LLMs can automate preliminary screening, plagiarism detection, and language checks, reducing reviewer workload and increasing consistency. Significant ethical concerns include potential algorithmic bias, lack of transparency, and risks to data privacy and confidentiality. LLMs should supplement, not replace, human judgment, as human reviewers are essential for contextual understanding and nuance. There is a pressing need for clear, well-disseminated policies and ethical guidelines for using LLMs in peer review | Extend the sample size to enhance generalizability. Use mixed-methods approaches for deeper insights. Explore the development of more advanced AI tools tailored to aid the review process while addressing ethical concerns. Investigate the impact of LLMs on the workload and evolving role of human reviewers |
| Erel et al., 2025 | AI detectors: Copyleaks and ZeroGPT. The study investigates the use of generative AI writing tools (like ChatGPT, Bing, Bard) indirectly via these detectors | A high percentage (83.2%) of abstracts were classified as «AI-assisted». AI assistance was significantly higher among non-native English-speaking authors and in articles published in high-impact (SCI/SCIE) journals. A positive correlation was found between abstract length and the detected level of AI assistance | To incorporate a broader range of journals beyond the Web of Science and to scan complete manuscripts instead of only abstracts for a more comprehensive analysis |
| Erol et al., 2025 | Text Generators: ChatGPT 3.5, ChatGPT 4, ChatGPT 4o. AI-Output Detectors: GPTZero, ZeroGPT, Corrector App. Plagiarism Detector: Plagiarism Detector tool | AI-output detectors (GPTZero, ZeroGPT, Corrector) showed moderate to high success (AUC 0.75-1.00) in distinguishing AI-generated texts, with performance improving against newer ChatGPT versions. AI-generated content had very high originality scores (approx. 99%), evading plagiarism detection. No detector achieved 100% reliability, and false positives occurred, misclassifying up to 30.4% of human-written texts as AI-generated | Evaluate outputs from other large language models beyond ChatGPT. Incorporate factual accuracy assessment of AI-generated content. Develop transparent and interpretable detection models tailored to scientific writing. Investigate the impact of paraphrasing on detection evasion |

| Reference | Type of GenAI / AI-Powered Tools | Key Findings | Further Research Agenda |
|---|---|---|---|
| Fakharifar et al., 2025 | ChatGPT (GPT-4, GPT-4o), Claude, Google Gemini, Microsoft Copilot (Bing) | AI tools can enhance efficiency in drafting, literature synthesis, and content creation but require human oversight for accuracy and reliability. Performance varies between models, with ChatGPT-4 showing higher quantitative accuracy. Significant challenges include potential for errors, inherent biases in training data, and ethical concerns like plagiarism. The 2024 ICMJE guidelines mandate transparency, disclosure, and author accountability for AI-assisted content. Major journals have adopted cautious policies, often prohibiting AI in confidential peer review processes | Comprehensive validation of AI applications using diverse datasets. Development of standardized evaluation protocols. Longitudinal studies to assess long-term effects. Further investigation into bias detection and mitigation strategies. Comparative analysis of other AI tools (e.g., Claude, Gemini, Copilot) against ChatGPT |
| Fung & Ng, 2025 | Large Language Models (LLMs) and AI code generation tools, specifically mentioned: Poe (by Quora). Other tools referenced in a broader context include OpenAI ChatGPT and Microsoft Copilot | AI drastically lowers the technical and resource barriers (time, cost, expertise) of traditional Simulation-Based Research (SBR) by automating code generation. A case study demonstrated that functional simulation code could be generated in 20 seconds, a task that traditionally takes months. AI-SBR offers advantages over traditional methods, including enhanced predictive power for «what-if» scenarios, massive scalability (e.g., simulating 100,000 individuals), and feasibility for studying logistically challenging topics. AI-SBR is positioned as a complementary research paradigm to traditional empirical and normative methods | Develop robust validation standards and protocols for AI-generated simulation code. Explore hybrid methodologies that combine AI-SBR with targeted empirical data collection. Integrate AI-driven simulation competencies into research training and curricula |
| Ghasemi et al., 2025 | ChatGPT (OpenAI), specifically versions 3.5 and 4 | Advantages: Clinical decision support, improved medical education, research assistance, improved accuracy, self-learning, summarization, enhanced critical thinking, interaction abilities.<br>Limitations: Knowledge limitations and accuracy, reliability issues, outdated or biased training data, clinical decision-making challenges, ethical and policy concerns, patient safety risks, wrong reference lists, education and monitoring needs, responsibility and decision risks. | Improve knowledge base, reduce hallucinations, enhance ability to handle uncertainty, validate performance in clinical settings, establish ethical guidelines, ensure patient safety, and integrate with human oversight |
| Goyibova et al., 2025 | ChatGPT, GPT-3, GPT-4, Grammarly, Jenni AI, Google Gemini, Quillbot, Typeset, Scite, Poe | Benefits: Enhances writing efficiency, provides personalized feedback, aids research, and automates tasks.<br>Ethical Challenges: Raises concerns about plagiarism ("Algiarism"), academic integrity, over-dependence, and originality loss.<br>Research Clusters: Five key clusters identified: AI integration, language models, academic context, assessment/integrity, and student experience. | Develop ethical guidelines and frameworks for AI use in academia.<br>Investigate AI's long-term impact on critical thinking and writing skills.<br>Explore AI detection tools and methods to ensure academic integrity.<br>Examine AI's role in personalized learning and self-regulated writing.<br>Study cross-cultural and cross-linguistic applications of AI writing tools |

| Reference | Type of GenAI / AI-Powered Tools | Key Findings | Further Research Agenda |
|---|---|---|---|
| Hosseini et al., 2025 | Generative AI (GenAI), Large Language Models (LLMs), ChatGPT, AI-assisted writing tools integrated into word processors | Position Shift: The authors have changed their position from supporting mandatory disclosure to advocating for voluntary disclosure of GenAI use for writing assistance.<br>Three Justifications Rebutted: 1) Demarcating human vs. AI contribution is impractical; 2) Detailed disclosure is too burdensome and doesn't aid evaluation; 3) Disclosure can introduce bias against non-native English speakers and compromise blind peer review.<br>Proposal: Implement a voluntary, machine-readable checkbox system for disclosure at submission, with information visible only after publication to prevent bias | Not explicitly stated |
| Hu et al., 2025 | ChatGPT (Generative AI, Large Language Model) | Identified five identity configurations: Reluctant Adoption, Conditional Alignment, Strategic Realignment, Lingering Dissonance, and Reflective Congruence.<br>Identity negotiation is mediated by the tension between self-image, ideal self, and self-esteem (Rogers' Theory).<br>Disciplinary differences exist: STEM/applied fields view AI as a pragmatic tool, while humanities/social sciences experience deeper identity conflicts.<br>ChatGPT acts as a mediating technology (Verbeek's Theory), amplifying competence but sometimes reducing perceived authorship agency.<br>Using AI for text generation causes more significant identity conflict than using it for language polishing. | Examine the influence of socio-cultural contexts, institutional expectations, and support systems.<br>Investigate the impact of demographic factors (e.g., writing proficiency, years of experience).<br>Conduct longitudinal studies to track evolving identity configurations.<br>Undertake deeper ethical inquiry into authorship, originality, and transparency. |
| Huang et al., 2025 | Generative AI (GAI), Large Language Models (LLMs), ChatGPT, and other unspecified GAI tools | A majority of ophthalmology journals have established GAI guidelines. There is a strong consensus that GAI cannot be an author and that human authors are fully responsible for its outputs. Disclosure of GAI use is universally required. Guidelines for AI-generated images are common, with most journals permitting their use under specific conditions. A significant lack of standardization in guidelines exists across and within publishers. | Long-term monitoring of guideline evolution is needed. The development of standardized, universally accepted interdisciplinary guidelines for GAI use in academia is recommended |
| Hwang et al., 2025 | Large Language Models (LLMs), GPT-4. The study compared personalized (via in-context learning) and non-personalized versions | Writers define authenticity through the source of content, the construction/expression of their authentic self, and writing outcomes. Writers believe practices like starting with a clear vision and acting as "content gatekeepers" help preserve authenticity. Writers subjectively preferred personalized AI but showed no significant behavioral difference in usage. Readers could not reliably distinguish AI-assisted work, were interested in it, and reacted positively to writers using AI, contrasting with writers' concerns about negative reader perceptions. | Investigating the distribution of work between humans and AI; Pacing and temporal adjustment of AI support; Developing tools to help communicate writers' internal states; Further exploring writer-reader relationships in the context of AI co-creation |

| Reference | Type of GenAI / AI-Powered Tools | Key Findings | Further Research Agenda |
|---|---|---|---|
| Hytten, 2025 | Generative AI, specifically the paid version of ChatGPT | The author experiences a deep ethical conflict when using AI, feeling it undermines the intellectual struggle integral to scholarship. AI use challenges traditional concepts of authorship and ownership of scholarly work. The nature of scholarly labor is changing, risking the loss of intellectual depth and rigor. Current disclosure norms are inadequate for capturing the extensive, collaborative nature of human-AI interaction. AI presents a uniquely seductive and potentially irresistible temptation for academics and students | How to preserve the meaning of "authorship" and "original scholarship" in an AI world. How to help students use AI ethically to support, not replace, cognitively challenging tasks. The long-term impact of AI on the capacity for informed judgment and the quality of academic arguments |
| Knight, 2025 | Not specified; the focus is on AI ethics guidelines broadly, not specific AI tools | There is no well-established model for developing ethics guidelines. Most guidelines include statements on their development approach and audience, but evidence use is largely at a global level rather than for specific contexts. Common structural features include overview statements and ethical principles, but key elements like cases and strategies are often missing. Very few guidelines address the specific intersection of AI and research ethics in education. Restrictive licensing often prevents the adaptation and reuse of guidelines | Developing and evaluating shared processes and structures for guideline creation. Investigating methods for including evidence-mobilization materials in guidelines. Evaluating guidelines as resources for learning ethical practice. Exploring the distinctive ethical features of using AI in specific domains like education |
| Kumar et al., 2025 | ChatGPT (a large language model by OpenAI) | ChatGPT significantly improved scores for understanding article structure, mastering writing techniques, and using correct language. It had no statistically significant impact on identifying research problems, knowledge of research types, or understanding research methods. AI-generated content improved efficiency but often lacked originality and critical depth | Explore longitudinal outcomes and long-term skill retention. Standardize AI usage protocols. Develop institutional guidelines for responsible AI use. Investigate AI's role in scientific publishing ethics and its influence on writing development |
| Lendvai, 2025 | ChatGPT (OpenAI's large language model) | The discourse is structured into ten major thematic clusters, grouped into three areas: Risks and ethical challenges (e.g., hallucinated references, plagiarism, authorship ambiguity), Practical applications (e.g., assisting non-native speakers, language enhancement), and Innovation and scholarly integration (e.g., pedagogical frameworks, stylistic features). There is a dual narrative recognizing ChatGPT's benefits for accessibility and efficiency alongside serious concerns about academic integrity and the limitations of AI-detection tools | Move beyond identifying risks to examine how journal editors and institutions are adapting. Explore ChatGPT's role in promoting or complicating scholarly equity for non-native speakers and under-resourced researchers. Develop stronger theoretical frameworks and conduct longitudinal studies on AI's impact on scholarly communication |

| Reference | Type of GenAI / AI-Powered Tools | Key Findings | Further Research Agenda |
|---|---|---|---|
| Lim et al., 2025 | Microsoft Copilot, ChatGPT (GPT-3.5), Google Gemini 1.0 | Generative AI tools were significantly more effective than human reviewers in providing feedback (mean score 3.44 vs. 2.33, $p<0.001$) and significantly more time-efficient (11.08 min vs. 45.15 min, $p<0.001$). Human reviewers performed better in providing recommendations (mean score 3.36 vs. 2.85, $p<0.01$). AI tools sometimes produced fictitious content ("hallucinations"). | Development of customizable AI assistants trained on specific datasets of peer-reviewed articles and academic writing styles; exploration of AI tools with fewer limitations (e.g., paid versions, graphic and table processing capabilities) |
| Liu et al., 2025 | ChatGPT | ChatGPT demonstrates significant potential in overcoming writing anxiety, improving writing efficiency, generating initial scientific papers, and assisting researchers and students with feedback. However, it faces challenges in data accuracy, ethics of generated content, authorship attribution, and risks of plagiarism and bias | Future studies should explore addressing challenges through technical and policy improvements, establishing ethical guidelines and legal frameworks, designing long-term tracking studies, optimizing algorithms for accuracy, and promoting multilingual interaction and collaboration |
| Miller et al., 2025 | ChatGPT; AI detection via ZeroGPT | A significant increase in suspected AI use in abstracts after ChatGPT's launch (21.1% in 2022 vs. 30.1% in 2024, $p=0.0002$). No significant increase was found in full-text articles. European publications had significantly lower AI content in abstracts (>10% threshold) compared to other regions | Investigate AI use in submitted vs. published manuscripts; study the impact of AI use on acceptance rates; use multiple detection tools and author surveys for enhanced validation; track emerging trends over time |
| Mo & Crosthweite, 2025 | ChatGPT-4.0, ERNIE Bot 4.0, Meta AI (using Llama 3.1) | Human academic writers use a significantly higher frequency and wider range of stance and engagement features compared to GenAI. LLMs use a narrower, more repetitive set of features. Disciplinary use was largely similar except in Philosophy, where AI underused reader pronouns and shared knowledge appeals. Significant variation was also found between the three LLMs' performance | Investigate stance and engagement in "hard" disciplines; use automated annotation on larger text samples; explore the impact of different prompting strategies; include other competitor LLMs (e.g., Claude, Gemini). |
| Mohan et al., 2025 | Large Language Models, ChatGPT, AI-assisted technology (mentioned as tools requiring acknowledgment in manuscripts) | The EIC's role is multifaceted, requiring management of peer review, plagiarism detection, ethical issues like guest/ghost authorship, salami slicing, and duplicate publications. Best practices include using software for screening, double-blinded peer review, incentivizing reviewers, and collaborating with other editors | Not explicitly stated |
| Mondal et al., 2025 | AI Text Generators (e.g., ChatGPT, Claude, Perplexity); AI Detection Tools (e.g., Turnitin); AI Paraphrasers/ Humanizers (e.g., Stealth Writer, QuillBot) | AI tools can assist in drafting, editing, and refining academic text, especially for non-native English speakers. However, over-reliance or misuse risks undermining originality, clarity, and intellectual contribution. AI detectors are often unreliable and can be bypassed by humanizer tools, creating a cycle that complicates academic integrity. Transparency, human oversight, and adherence to ethical guidelines (e.g., ICMJE, WAME) are essential for responsible AI use | Not explicitly stated |

| Reference | Type of GenAI / AI-Powered Tools | Key Findings | Further Research Agenda |
|---|---|---|---|
| Moorhouse et al., 2025 | Large Language Models (LLMs) (e.g., ChatGPT-4); AI-powered research tools (e.g., Open AI's Deep Research, Scopus AI); Computer-Assisted Qualitative Data Analysis Software (CAQDAS) with AI features (e.g., MAXQDA, NVivo); AI summarization and dissemination tools (e.g., Google NotebookLM) | A framework built on five elements of quality study:<br>(1) transparency,<br>(2) methodological rigor,<br>(3) ethics,<br>(4) societal value, and<br>(5) human accountability<br>It can guide TESOL researchers in making principled decisions about using GenAI. The framework provides methodological actions and reflective questions for researchers and can be used by journals to evaluate AI use | The framework requires further specialization for specific research methodologies and TESOL communities. The competencies needed to use GenAI for TESOL research need to be fully articulated. The framework should be tested on various study types beyond qualitative interview studies. |
| Nassar et al., 2025 | Large Language Models (LLMs) (e.g., ChatGPT); AI detection tools (e.g., ZeroGPT) | 3.28% of 3,596 articles published after ChatGPT's release exceeded the AI-generated content threshold. Use varied significantly by journal (Arthroscopy: 7.17%; KSSTA: 0.93%). A significant upward temporal trend was identified, with AI use increasing from 2.38% in early 2023 to 6.25% in late 2024. Editorial policies appear to influence usage rates | Future studies should validate findings using other AI detection tools. Research should assess the accuracy of AI-generated content and its impact on article integrity. Studies should evaluate the adequacy of AI usage declarations in manuscripts and investigate author compliance with journal policies |
| Öztürk et al., 2025 | Generative Pre-Trained Transformer 4 omni (GPT-4o); ChatGPT | 78.6% (11 of 14) of reviewers did not realize the manuscript was AI-generated. At the editorial stage, 14.3% accepted it directly and 42.9% sent it for peer review. As peer reviewers, 28.6% recommended acceptance after minor revisions. Only 21.4% correctly identified the AI origin, citing standardized structures, superficial discussion, and repetitive language as clues | Not explicitly stated |
| Parlak et al., 2025 | ChatGPT-4o (GPT-4 Omni) | AI-generated titles were significantly preferred (58.1%) over original human-written titles (41.9%). The most cited reason for preference was "comprehensibility" (47.3%). AI titles were favored for clarity and attractiveness, while original titles were more frequently chosen for "scientific accuracy." 93.3% of participants preferred AI titles more frequently overall | Future studies should compare outputs from different AI systems, involve larger participant samples, sample titles from multiple journals, and evaluate titles at a contextual level beyond just linguistic features. |
| Picardal et al., 2025 | ChatGPT (OpenAI), Perplexity AI, Gemini, Unriddle, CICI (Classroom Intelligent Companion Interface), Microsoft Copilot, Quillbot, Grammarly, Google Assistant, Scribbr | ChatGPT was the most used AI tool (62%). AI was primarily used for content generation (19%), data analysis/interpretation (17%), and methodological guidance (14%). Students employed diverse prompting strategies (direct questioning, contextual, iterative, role-based). AI served as a cognitive and metacognitive scaffold, but concerns emerged regarding over-reliance, academic integrity, and the need for critical oversight. Seven contextual factors influenced AI use, including cognitive load, time pressure, and lack of confidence in writing | Longitudinal studies to trace how AI-literate preservice teachers integrate AI into their professional practice post-graduation; investigating the impact of AI-supported writing on deeper learning outcomes; examining the role of smart prompting as a 21st-century literacy skill and its implications for curriculum design and assessment across disciplines. |

| Reference | Type of GenAI / AI-Powered Tools | Key Findings | Further Research Agenda |
|---|---|---|---|
| Prakash et al., 2025 | Large Language Models (LLMs) such as ChatGPT, AI-assisted writing tools like Grammarly | Global improvement in writing complexity from 2012 to 2024, with notable gains in non-native English-speaking and lower-income countries; China surpassed the United States in readability scores; Internet penetration is a significant predictor of writing quality; Increased use of adjectives and adverbs associated with LLM-generated text, especially in non-native English-speaking countries and China, suggesting wider adoption of AI tools | Examine AI's broader influence on education, creativity, and communication; Investigate the impact of AI tools on foundational writing skills and academic integrity in primary/secondary education; Explore AI's role in creative outputs and professional communication; Use qualitative assessments to evaluate deeper writing qualities beyond surface-level readability; Conduct original analyses to identify evolving linguistic markers of AI-generated content |
| Premat et al., 2025 | Generative AI tools, specifically AI chatbots like ChatGPT | Swedish HE approach emphasizes trust, pedagogical support, and context-sensitive reasoning over surveillance; 57% of students support disclosing ChatGPT use, especially for substantial content generation; Students exhibit a pragmatic, context-sensitive approach, distinguishing between substantial and auxiliary AI use; Institutional guidelines evolved from initial prohibitions towards promoting transparency and responsible use | Explore similar questions across diverse disciplines, universities, and cultural contexts; Combine student, teacher, and policy-maker perspectives; Conduct longitudinal studies to trace the evolution of student attitudes alongside institutional policies and AI technologies |
| Pretorius et al., 2025 | Generative AI tools, specifically ChatGPT; also mentions QuillBot, Grammarly, and Canva | GenAI reshaped power dynamics and challenged academic hierarchies by providing real-time language support; It empowered multilingual scholars to participate in academic discourse, disrupting traditional gatekeeping; By automating routine tasks, GenAI shifted the academic focus from technical skills to intellectual contributions; GenAI supported personalised learning, socialisation into academic communities, and wellbeing; The ethical use of GenAI, framed by Ubuntu, acts as a democratising force promoting social justice and inclusivity | Not explicitly outlined |
| Radtke & Rummel, 2025 | AI-based text generators (e.g., ChatGPT, GPT-3.5, Jasper, Google Gemini/Bard); Large Language Models (LLMs) | Learners invested less time revising a text labeled as AI-generated compared to one labeled as peer-written, but the number of identified areas for improvement and revision operations did not differ significantly. Greater prior exposure to media reports about AI, higher trust in AI, and a tendency toward social loafing in AI-assisted writing were associated with less revision time for AI-labeled texts. Learners with more experience in academic writing identified more areas for improvement and made more revisions, regardless of authorship. Female learners revised more extensively than male learners across all measures, irrespective of the labeled author | Investigate the underlying psychological processes of revising peer-written vs. AI-generated texts; Examine revision focus (e.g., factual accuracy for AI texts vs. linguistic issues for peer texts) and depth (surface vs. meaning-changing revisions); Explore the role of learner-AI interaction prior to revision; Include other authorship labels (e.g., expert, peers with varying expertise); Use think-aloud protocols and alternative revision measures (e.g., characters changed, text length); Investigate the role of gender-related self-concept beliefs and technology self-efficacy in AI-assisted writing; Develop improved measures for writing skills and AI literacy |

| Reference | Type of GenAI / AI-Powered Tools | Key Findings | Further Research Agenda |
|---|---|---|---|
| Rafi & Amjad, 2025 | Generative AI (GenAI), ChatGPT, GPT-4, AI writing bots, AI simulation tools, AI sentiment analysis tools | GenAI aids in improving writing mechanics, providing efficient feedback, augmenting thinking processes, structuring arguments, and facilitating data analysis. Challenges include hallucinations, ethical concerns, context-text mismatch, and reduced critical thinking. Strategies include supervisor guidance, recognizing AI linguistic cues, and effective prompt formulation | Experimental studies to quantify GenAI's impact on cognitive development; correlation between prompt formulation skills and GenAI output; changes in supervisor-student interaction dynamics due to GenAI reliance |
| Ragel, 2025 | AI and Generative AI (GenAI) technologies; mentioned tools: Grammarly | AI and GenAI are transformative tools for scholarly publishing, aiding in data analysis, literature reviews, and drafting. They help remove linguistic barriers, accelerate discoveries in fields like genomics and climate science, reduce cognitive biases, and promote more reproducible research. Transparency, ethical guidelines, and equitable access are crucial for responsible use | Not explicitly stated |
| Raitskaya & Tikhonova, 2025 | Generative AI (GenAI), ChatGPT, large language models (LLMs) such as BERT and LLM4HA | GenAI holds substantial potential to support critical thinking development, particularly when pedagogically integrated. 21 out of 30 reviewed studies reported a positive impact. Key benefits include support for idea generation, argument construction, and metacognitive awareness. Risks include cognitive offloading and over-reliance. Impact is moderated by individual factors (e.g., motivation, AI literacy) and pedagogical design | Longitudinal studies on long-term effects; discipline-specific instructional models; robust theoretical frameworks linking AI use to cognitive outcomes; conditions under which GenAI fosters meaningful learning |
| Resnik & Hosseini, 2025 | AI tools, generative AI, AI-assisted technologies, chatbots, large language models | Proposes a framework where AI disclosure should be mandatory only when use is intentional and substantial. Substantial use is defined by three criteria: 1) AI makes decisions that directly affect research results, 2) AI generates/synthesizes content/data/images, or 3) AI analyzes content/data/images. Disclosure is optional for minor assistance (e.g., editing) and unnecessary for trivial/integrated uses | Survey journal editors and experts from different scholarly domains to understand their views on disclosure policies; further study on whether disclosing optional AI use to peer reviewers introduces bias. |
| Reza et al., 2025 | Large Language Models (LLMs); AI-assisted writing tools (e.g., ChatGPT, Grammarly, Microsoft Copilot, Claude, LLaMA, Sudowrite); Systems for brainstorming, drafting, editing, and providing feedback | Four primary design strategies were identified in HCI research: Structured Guidance, Guided Exploration, Active Co-Writing, and Critical Feedback. Writers' ownership preferences vary by writing context. Content-centric writers prioritize control over planning, while form-centric writers prioritize control over translating and reviewing. Writers' sense of ownership is modulated by contextual factors like time constraints, task importance, confidence, and trust. Interaction design is crucial, with writers preferring AI suggestions over direct edits, maintaining final say, and having global and local AI toggles. The cognitive process of Monitoring is critically underexplored in AI-assisted writing research | Develop more flexible systems that allow users to adjust AI involvement across different writing processes. Investigate and design support for the underexplored Monitoring cognitive process. Incorporate global and local AI toggles in interaction design to minimize distractions and grant fine-grained control. Expand research to include more age-diverse, educationally diverse, and AI-naive participant samples. Explore alternative theoretical frameworks beyond Flower and Hayes to better describe Human-AI collaboration. |

| Reference | Type of GenAI / AI-Powered Tools | Key Findings | Further Research Agenda |
|---|---|---|---|
| Sparkman & Witt, 2025 | Claude (Anthropic); Large Language Models (LLMs); Generative AI; ChatGPT is also discussed extensively in the literature and the review content | Claude excelled at summarizing points of agreement across sources and producing coherent, introductory-style text. Key limitations included providing incomplete citations, hallucinating false information (e.g., incorrect page numbers), struggling with synthesis, and producing mistargeted citations that misrepresented source arguments. The AI was unable to generate original thought or analysis beyond the provided articles and could not integrate outside knowledge or lived experience | Repeating the experiment using the paid version of Claude to overcome the technical limitations of the free version. Conducting repeated testing on various and newer LLM models to track their progress in performing scholarly research tasks |
| Suchikova et al., 2025 | Generative AI (GAI) / Large Language Models (LLMs); specifically mentioned tools include ChatGPT, Claude, DeepSeek, and Gemini | Introduces the GAIDeT taxonomy, which classifies GAI contributions across macro and micro levels of the research process (e.g., Conceptualization, Literature Review, Methodology). A web-based GAIDeT Declaration Generator was developed to help researchers create transparent disclosure statements. The framework maintains that GAI cannot be an author but its contributions must be documented to preserve human accountability | Future work should focus on empirical validation, cross-disciplinary adaptability, policy implications for GAI governance, and developing standardized methods for determining proportional GAI credits |
| Veiga et al., 2025 | Generative AI (e.g., ChatGPT, other Large Language Models); AI-assisted tools (e.g., Grammarly, Mendeley, EndNote, Zotero). | Inconsistencies exist among top publishers' AI policies. AI-assisted tools are generally accepted for grammar and formatting, while generative AI use requires caution and disclosure. All publishers prohibit AI authorship. Key themes are author accountability, human oversight, and transparency. Concerns include biases, quality/reliability issues, and intellectual property compliance | Not explicitly stated |
| Yaqin et al., 2025 | ChatGPT (a Generative AI / Large Language Model developed by OpenAI) | The Untrained Dataset (UTD) had significantly more grammatical errors (87) than the Trained Dataset (TD) (18). Common errors in UTD included incorrect sentence structure, omission of nouns, and inappropriate word choices. Structured prompts improved output quality but did not eliminate persistent errors like phrase structure and affixation issues | Future research should focus on fine-tuning AI models with syntactic rules, expanding Malay linguistic corpora, conducting comparative studies with other low-resource languages, and investigating the effectiveness of AI-assisted writing tools in academic settings |
| Yin & Chapelle, 2025 | Generative AI (GenAI) / Large Language Models (LLMs); specifically mentioned tools include ChatGPT, Gemini, Midjourney, DALL-E, and Sora | Less than half (44.7%) of the 170 applied linguistics journals provided GenAI use guidelines. There was strong consensus against AI authorship (96.1%) but significant disagreement on other uses (e.g., citing AI, manuscript drafting, image generation). Guidelines were often incomplete, with low coverage for data-related tasks. Inconsistencies were found both across journals and between journals and their publishers | Future research should monitor how AI policies evolve over time and investigate a broader range of journals and publishers. The study also calls for the applied linguistics community to address how AI guidelines are formulated and to find ways to involve more people in creating standardized guidance |

| Reference | Type of GenAI / AI-Powered Tools | Key Findings | Further Research Agenda |
|---|---|---|---|
| Yoo, 2025 | Generative AI, Large Language Models, ChatGPT, Gemini, AI-assisted writing tools, AI detection tools (e.g., GPTZero, Originality. AI, Turnitin AI Detector) | International consensus prohibits AI authorship; transparent disclosure of AI use is mandatory; human authors retain full accountability; AI detection tools are unreliable and biased; journal policies vary in strictness but share core ethical principles | Developing more nuanced disclosure standards; creating educational programs for stakeholders; establishing guidelines for AI use in peer review; preparing for future scenarios where AI reaches human-level writing proficiency |

# APPENDIX 4

## Uses of Generative AI in scholarly writing and publishing

| Use Category | Specific Use | References |
|---|---|---|
| Idea Generation & Conceptualization | Brainstorming, generating hypotheses, outlining | Akgun et al., 2025; Alamri et al., 2025; Archana et al., 2025; Bobier et al., 2025; Cohen & Moher, 2025; Hosseini et al., 2025; Lendvai, 2025; Picardal, 2025; Premat & Farazouli, 2025; Resnik & Hosseini, 2025; Sparkman & Witt, 2025; Yoo, 2025 |
| Literature Review & Sourcing | Literature searching, summarizing, gap identification | Alamri et al., 2025; Archana et al., 2025; Fakharifar et al., 2025; Lendvai, 2025; Resnik & Hosseini, 2025; Sparkman & Witt, 2025; Suchikova et al., 2025 |
| Writing & Drafting | Drafting text, paraphrasing, synthesizing | Akgun et al., 2025; Alamri et al., 2025; Archana et al., 2025; Bobier et al., 2025; Cohen & Moher, 2025; Erol et al., 2025; Fakharifar et al., 2025; Ghasemi et al., 2025; Lendvai, 2025; Picardal, 2025; Resnik & Hosseini, 2025; Yoo, 2025 |
| Editing & Language Enhancement | Grammar checking, improving readability, polishing | Akgun et al., 2025; Al Hosni, 2025; Archana et al., 2025; Cohen & Moher, 2025; Goyibova et al., 2025; Hosseini et al., 2025; Lendvai, 2025; Lim G.H. et al., 2025; Nassar et al., 2025; Prakash et al., 2025; Premat & Farazouli, 2025; Resnik & Hosseini, 2025; Yoo, 2025 |
| Data-Related Tasks | Data analysis, generation, collection, coding | Alamri et al., 2025; Archana et al., 2025; Fakharifar et al., 2025; Fung & Ng, 2025; Resnik & Hosseini, 2025; Suchikova et al., 2025; Yoo, 2025 |
| Visualization & Illustration | Generating figures, tables, images | Resnik & Hosseini, 2025; Yoo, 2025 |
| Research Design & Methodology | Study design, methodology, simulation coding | Fung & Ng, 2025; Picardal, 2025; Resnik & Hosseini, 2025; Suchikova et al., 2025 |
| Peer Review & Feedback | Providing feedback on drafts, assisting review | Ebadi et al., 2025; Lim G.H. et al., 2025 |
| Administrative & Support Tasks | Formatting references, managing citations | Archana et al., 2025; Resnik & Hosseini, 2025; Yoo, 2025 |

# APPENDIX 5

## Themes of Publications included in the Review

| Nos | Reference | Publication Theme | Authors' Keywords | Coded Theme |
|---|---|---|---|---|
| 1 | Akgun et al., 2025 | Comparing the quality of neurosurgery articles written by human authors and those generated by AI (ChatGPT) | ChatGPT, GPT-4, Artificial Intelligence, article | 2, 3, 5 |
| 2 | Al Hosni, 2025 | Preserving authorial voice and linguistic fingerprint in academic writing in the context of Generative AI use, particularly for second language learners | Academic writing, academic integrity, authorial voice, ChatGPT, generative artificial intelligence, linguistic fingerprint, student identity, stylistic analysis | 2, 4, 6, 9, 10 |
| 3 | Alamri et al., 2025 | A SWOT analysis of the use of AI, specifically ChatGPT, in academic writing | Artificial Intelligence (AI), academic writing, ChatGPT, SWOT Analysis Framework | 4 |
| 4 | Alghazo et al., 2025 | Engagement strategies in human-written and AI-generated academic essays | Engagement strategies, academic writing, human-written, AI-Generated, appraisal theory | 4 |
| 5 | Alkhawam et al., 2025 | AI use policies in manuscript writing in cardiology and vascular journals | Artificial intelligence, AI, large language models, LLM, cardiology journals, vascular journals, AI policies, ethical standards, journal policy | 1 |
| 6 | Alnaimat et al., 2025 | Perspectives of Artificial Intelligence Use for In-House Ethics Checks of Journal Submissions | Artificial Intelligence, academic publishing, ethics, peer review, authorship | 1, 7 |
| 7 | Anghelescu et al., 2025 | Detection, minimization, and elimination of errors (hallucinations) in Large Language Models (LLMs) and enhancing their reliability, particularly in biomedical domains | Artificial intelligence (AI), large language models (LLMs), ChatGPT, RAG (Retrieval-Augmented Generation systems), ACURAI | 3, 5 |
| 8 | Arar et al., 2025 | The ethical implications, research integrity, and academic use of Artificial Intelligence (AI) and Generative AI (GenAI) | Artificial intelligence, generative artificial intelligence, research integrity, ethics, bibliometric analysis, systematic review | 2, 8 |
| 9 | Archana et al., 2025 | AI-assisted learning and research among university students and scholars, focusing on adoption, usage patterns, and influencing factors | AI adoption, GenAI tools, AI tools, ChatGPT adoption, AI for research, AI-assisted learning, Binary logistic regression | 4, 6 |
| 10 | Ateriya et al., 2025 | The ethical implications, challenges, and proposed guidelines for using Artificial Intelligence (AI) in academic writing and research publication | Artificial intelligence, academic writing, authorship attribution, plagiarism detection, bias mitigation | 2 |
| 11 | Bobier et al., 2025 | Artificial intelligence policies in bioethics and health humanities journals | Research ethics, bioethics scholarship, artificial intelligence, academic publishers | 1 |
| 12 | Cambraia & Pyrrho, 2025 | The risk of technological colonialism posed by Generative Artificial Intelligence | Artificial intelligence, generative artificial intelligence, large language models, digital ethics, digital colonialism | 2, 10 |
| 13 | Cohen & Moher, 2025 | The opportunities and threats of using generative AI in medical and academic writing | Artificial Intelligence; large language models; medical writing; publication ethics; research ethics; research integrity | 1, 5 |

| Nos | Reference | Publication Theme | Authors' Keywords | Coded Theme |
|---|---|---|---|---|
| 14 | Comas-Forgas et al., 2025 | The linguistic impact of Generative AI on academic article titles, specifically the analysis of verbs suspicious of being AI-generated/ assisted | Artificial intelligence, databases, scientific communication, titles | 4 |
| 15 | Daoudi, 2025 | Ethical limits and suggestions for improving the use of AI in scientific research, academic publishing, and the peer review process | Artificial Intelligence, ethical boundaries, scientific research, academic publishing, peer review | 2 |
| 16 | Ebadi et al., 2025 | The impact of Generative AI (specifically Large Language Models) on the academic peer review process, focusing on the perspectives and experiences of journal reviewers | Large language models (LLMs), peer review, journal review, academic publishing, reviewer perspectives | 5, 7 |
| 17 | Erel et al., 2025 | Examining the frequency and correlates of AI-generated content in abstracts from anaesthesiology and intensive care journals | Anaesthesiology, artificial intelligence, artificial intelligence detector, journal, technology | 2, 3 |
| 18 | Erol et al., 2025 | Evaluating the reliability and accuracy of AI-output detection tools in distinguishing human-authored academic content from texts generated by different versions of ChatGPT | Academic writing, AI-output detector, artificial intelligence, ChatGPT, ethical considerations, plagiarism detection | 3 |
| 19 | Fakharifar et al., 2025 | The applications, benefits, limitations, and ethical considerations of Artificial Intelligence (AI) and Large Language Models (LLMs) like ChatGPT in medical writing | Medical writing, Chat Generative Pre-trained Transformer (ChatGPT), large language models (LLMs), ethics, International Committee of Medical Journal Editors guidelines (ICMJE guidelines) | 1, 5 |
| 20 | Fung & Ng, 2025 | The potential of Artificial Intelligence (AI) to overcome traditional barriers and enable a new methodology known as AI-driven Simulation-Based Research (AI-SBR) | Artificial Intelligence (AI), simulation-based research, empirical research, methodology, normative research, traditional research, research method | 5 |
| 21 | Ghasemi et al., 2025 | Advantages and limitations of ChatGPT in healthcare | Advantages, ChatGPT, health, limitations, scoping review | 5 |
| 21 | Goyibova et al., 2025 | The impact of Artificial Intelligence on academic writing, including trends, advancements, and ethical challenges | Artificial Intelligence (AI), academic writing, impact, bibliometric, ChatGPT, grammar | 2, 4, 8 |
| 22 | Hosseini et al., 2025 | The ethics and policy of disclosing the use of Generative AI for writing assistance in scholarly manuscripts | Artificial intelligence, disclosure, editorial policies, publication ethics, peer review, writing | 1, 7 |
| 23 | Hu et al., 2025 | Identity negotiation and construction of non-native English researchers when using ChatGPT for academic writing | Not explicitly listed | 4, 6, 9, 10 |
| 24 | Huang et al., 2025 | Guidelines on the use of Generative Artificial Intelligence (GAI) in ophthalmology journals | Not explicitly listed | 1 |
| 25 | Hwang et al., 2025 | Authenticity in human-AI co-writing, specifically how writers and readers perceive and seek to preserve the writer's authentic voice when using Large Language Models (LLMs) | Generative AI, AI-assisted writing, Large Language Model, authenticity, creativity, Human-AI Collaboration | 3, 4, 9 |
| 26 | Hytten, 2025 | The ethical complexities of using generative AI in academic writing, focusing on authorship, scholarly labor, and disclosure | Ethics, generative AI, inquiry, academic labor | 2, 9, 10 |

| Nos | Reference | Publication Theme | Authors' Keywords | Coded Theme |
|---|---|---|---|---|
| 27 | Knight, 2025 | AI ethics guidelines development, structure, and use of evidence, with a focus on education and AI | Ethics guidelines, research ethics, AI ethics, scoping review, ethics review | 1, 6 |
| 28 | Kumar et al., 2025 | The impact of ChatGPT on the scientific writing skills of orthopedic residents | Artificial intelligence, scientific writing, medical education, natural language processing | 4, 5, 6 |
| 29 | Lendvai, 2025 | A scientometric analysis of the scholarly discourse on the role of ChatGPT in academic writing, mapping its thematic and intellectual structure | ChatGPT, academic writing, AI and research, ethical challenges, scholarly innovation | 8 |
| 30 | Lim H.G. et al., 2025 | Use of generative AI in the peer review process for occupational health manuscripts | Not explicitly provided | 1, 5, 7 |
| 31 | Liu et al., 2025 | Applications, limitations, and ethical challenges of ChatGPT in academic writing | ChatGPT; writing, authorship, ethics, data accuracy, plagiarism, students, bias | 2, 4 |
| 32 | Miller et al., 2025 | Prevalence and geographic correlation of AI-generated text in a shoulder and elbow surgery journal before and after ChatGPT's release | Artificial intelligence, shoulder, elbow, ChatGPT, AI-generated content, plagiarism detection, academic integrity | 2, 3, 5 |
| 33 | Mo & Crosthweite, 2025 | Comparison of stance and engagement features in academic writing produced by humans versus generative AI large language models | Stance; engagement, English for academic purposes, Generative AI, large language model, academic writing | 3, 4 |
| 34 | Mohan et al., 2025 | Challenges and best practices for Editors-in-Chief in maintaining quality, integrity, and ethical standards in radiology journal publishing | Editorial challenges, editorial gauntlet, radiology publishing | 1, 5, 7 |
| 35 | Mondal et al., 2025 | The ethical and practical challenges of using AI in academic writing, focusing on the balance between efficiency and integrity, and the interplay between AI text generators, detectors, and humanizers | Artificial intelligence, ChatGPT, large language model, scientific writing | 2, 3, 9 |
| 36 | Moorhouse et al., 2025 | Proposing a disciplinary framework for the responsible use of Generative AI (GenAI) in TESOL research to maintain and enhance study quality | Not explicitly provided | 1, 6 |
| 37 | Nassar et al., 2025 | Quantifying the prevalence and temporal trend of AI-generated content in published sports medicine research articles | Artificial intelligence, sports medicine, manuscript writing, research integrity | 2, 3, 5 |
| 38 | Öztürk et al., 2025 | Investigating whether academic journal editors and peer reviewers can detect a scientific manuscript entirely generated by an AI model | ChatGPT, artificial intelligence, journal, editor, reviewer | 2, 3, 7 |
| 39 | Parlak et al., 2025 | Comparing expert preferences and perceptions of human-written versus AI-generated article titles in Physical Medicine and Rehabilitation | Academic writing, artificial intelligence, physical medicine and rehabilitation | 2, 3, 5 |
| 40 | Picardal et al., 2025 | Investigating how preservice science teachers utilize AI-driven smart prompts to support academic research writing within a digitally mediated learning environment, grounded in the TPACK framework | Smart prompts, preservice teachers, academic writing, TPACK, AI in Education | 4, 6 |

| Nos | Reference | Publication Theme | Authors' Keywords | Coded Theme |
|---|---|---|---|---|
| 41 | Prakash et al., 2025 | The evolution of academic writing quality in social sciences abstracts, focusing on disparities across linguistic, regional, economic, and gender-based classifications, and the role of AI tools and digital infrastructure in mitigating these disparities | Not explicitly provided | 4, 5, 10 |
| 42 | Premat et al., 2025 | Conceptualizations and practices of academic integrity in Swedish higher education in the context of Generative AI tools, analyzing institutional responses and student perspectives | Academic integrity, Generative AI. ChatGPT, Swedish universities, institutional policy, phenomenography, ethics in higher education | 2, 6, 10 |
| 43 | Pretorius et al., 2025 | The use of Generative AI to empower international PhD students, decolonise academic communication, and promote epistemic justice, viewed through the theoretical lens of Ubuntu | Generative AI, AI Literacy, decolonisation, academia, doctoral education, ubuntu, autoethnography | 3, 6, 9, 10 |
| 44 | Radtke & Rummel, 2025 | The impact of authorship information (peer vs. AI) on learners' revision behavior in academic writing, including the influence of prior experiences, attitudes, and gender | AI-assisted writing, text revision, Generative AI, academic writing, collaborative writing | 4, 6, 9 |
| 45 | Rafi & Amjad, 2025 | The role of generative AI in doctoral dissertation writing, focusing on opportunities, challenges, and strategies to promote human agency | GenAI-assisted dissertation writing, GenAI-assisted writing opportunities and challenges, AI literacy program, GenAI and human interaction, Technology acceptance | 4, 6, 9 |
| 46 | Ragel, 2025 | The integration of Artificial Intelligence (AI) and Generative AI (GenAI) in scholarly publishing to enhance integrity, accessibility, and the research process | Not provided | 5 |
| 47 | Raitskaya & Tikhonova, 2025 | The impact of Generative AI (GenAI), particularly ChatGPT, on the development of critical thinking skills among university students | Development of critical thinking skills, generative AI impact on higher education, ChatGPT and student cognition, AI-supported academic writing, metacognitive engagement with GenAI, GenAI and argumentation, prompt engineering and critical thinking | 4, 6, 10 |
| 48 | Resnik & Hosseini, 2025 | Ethical framework for disclosing the use of Artificial Intelligence (AI) in scientific research and publication | AI, ethics, disclosure, transparency, accountability | 1, 2 |
| 49 | Reza et al., 2025 | Aligning the design of Human-AI collaborative writing tools with user demands to preserve writers' sense of agency and ownership across different writing processes and contexts | AI-Assisted Writing, Human-Centered AI, Large Language Models | 4, 6, 9 |
| 50 | Sparkman & Witt, 2025 | An experimental comparison of a human-written and an AI-generated literature review to evaluate the utility, capabilities, limitations, and ethical implications of using generative AI for academic research | Generative AI, ChatGPT, Claude, literature review, academic writing, LLMs | 4, 5 |

| Nos | Reference | Publication Theme | Authors' Keywords | Coded Theme |
|---|---|---|---|---|
| 51 | Suchikova et al., 2025 | Development of a taxonomy (GAI-DeT) for the transparent delegation of research tasks to Generative AI, focusing on accountability and integrity in scientific research and publishing | Human-AI collaboration, large language models, research workflow, taxonomy | 2, 9 |
| 52 | Veiga et al., 2025 | Analysis of publishers' policies to propose ethical guidelines for using Generative AI and AI-assisted tools in scholarly publishing, focusing on accountability, transparency, and distinguishing between accepted and restricted uses | Artificial intelligence, authorship, organizational policy, scholarly communication, generative artificial intelligence | 1, 2 |
| 53 | Yaqin et al., 2025 | Evaluating the performance and linguistic accuracy of ChatGPT in generating Malay academic texts, focusing on grammatical errors and the impact of prompt engineering | AI-generated text, ChatGPT, expert correction, linguistic accuracy, Malay academic writing | 3, 4 |
| 54 | Yin & Chapelle, 2025 | A systematic analysis of the scope, nature, and consistency of GenAI use guidelines provided by high-impact applied linguistics journals, leading to the conceptualization of GenAI literacy for research | GenAI literacy for research; GenAI use guidelines; Applied linguistics journals; scholarly publishing | 1, 6 |
| 55 | Yoo, 2025 | A comparative review of editorial policies on the ethical use of generative AI in scientific writing and publishing, with a focus on authorship, disclosure, and accountability | Generative AI, ChatGPT, authorship, writing, publishing | 1, 2 |

# APPENDIX 6

## Toolkit for AI Use in Scholarly Writing and Publishing: A Taxonomy and Guideline Framework

This toolkit synthesizes recommendations from 56 scholarly articles and publisher policies to provide a structured approach for authors, reviewers, institutions, and publishers.

## Part 1: The GAIDeT-Inspired Taxonomy of AI Use

This taxonomy categorizes AI use to help determine the appropriate level of oversight and disclosure. It adapts and consolidates frameworks from Resnik & Hosseini (2025) and Suchikova et al. (2025).

## A. Substantive / High-Stakes Use

*(Mandatory Disclosure & Rigorous Human Oversight Required)*
These uses involve AI making or significantly influencing core intellectual contributions.

***1. Conceptualization & Design:***

Generating research questions or hypotheses.
Assisting in study design or methodology development.
Identifying theoretical frameworks.

***2. Content Generation & Synthesis:***

Drafting entire sections of a manuscript (Introduction, Methods, Results, Discussion).
Paraphrasing or summarizing literature to create new synthesized text.
Generating abstracts from scratch.

***3. Data Analysis & Interpretation:***

Conducting statistical analysis or coding for data processing.
Interpreting data trends or results.
Generating synthetic data sets.

***4. Visual Content Creation:***

Creating or modifying figures, tables, or images for publication.

## B. Supportive / Medium-Stakes Use

*(Optional Disclosure & Active Human Oversight Recommended)*
These uses leverage AI for efficiency in essential tasks where the human author retains primary control.

***1. Writing Assistance:***

Restructuring or reorganizing human-written text for clarity.
Generating an initial outline based on author-provided points.
Brainstorming ideas or suggesting titles.

***2. Research Support:***

Conducting initial literature searches.
Summarizing existing articles (without synthesis).
Identifying relevant references (to be verified by the author).

***3. Analytical Feedback:***

Providing feedback on arguments or logic.
Validating code or analysis plans.

## C. Routine / Low-Stakes Use

*(Disclosure Typically Unnecessary & Basic Verification Sufficient)*
These uses involve AI as a integrated tool for polishing and administrative tasks, with minimal creative or intellectual input.

***1. Language Enhancement:***

Correcting grammar, spelling, and punctuation.
Improving readability and style (e.g., using tools like Grammarly).
Rephrasing sentences for fluency.

***2. Administrative Tasks:***

Formatting references and citations.
Checking journal submission guidelines.
Translating text (with human verification for accuracy).

## Part 2: Consolidated Ethical Guidelines & Best Practices

These guidelines are organized by stakeholder responsibility and core ethical principles.

***Core Principles for All Stakeholders***

1. **Human Accountability is Non-Negotiable:** AI is a tool, not an agent. Human authors bear ultimate responsibility for the accuracy, integrity, and entirety of the submitted work, including all AI-generated content. (ICMJE, COPE, WAME, et al.)
2. **Transparency Through Proportional Disclosure:** Disclose AI use in a dedicated section of the manuscript (e.g., "Declaration of AI Use"). The level of detail should be proportional to the "stakes" of the use, as defined in the taxonomy above.
3. **Preservation of Intellectual Labor:** AI should **augment, not replace**, critical thinking, deep analysis, and the development of a unique authorial voice. Guard against the erosion of scholarly skills and epistemic diversity.

## Guidelines for Researchers & Authors

| Ethical Imperative | Specific Actionable Guidelines |
|---|---|
| Authorship & Attribution | • Never list an AI as a co-author.<br>• Clearly define and claim your own intellectual contributions. |
| Transparency & Disclosure | • Disclose all "Substantive" and "Supportive" AI uses (see Taxonomy[3]).<br>• In the disclosure, specify: Tool Name & Version, Purpose of Use, and Extent of Contribution.<br>• When in doubt, disclose. |
| Integrity & Plagiarism | • Do not present AI-generated text as your own original work ("Algiarism").<br>• Verify all outputs for factual accuracy, plagiarism, and "hallucinations."<br>• Do not cite AI as an authoritative source unless journal policy explicitly permits it. |
| Critical Thinking & Voice | • Use AI as a "reasoning partner" or scaffold, not an answer generator.<br>• Actively revise and personalize all AI-generated content to preserve your authorial voice.<br>• Use AI for brainstorming but ensure core ideas and arguments are your own. |

[3] Suchikova et al., 2025

| | |
|---|---|
| Data Privacy & Security | • Never input confidential, unpublished, or sensitive data (e.g., patient information, peer-review materials) into public AI systems.<br>• Use enterprise-level AI tools with data privacy guarantees when possible. |
| Bias & Equity | • Critically audit AI outputs for demographic, cultural, and linguistic biases.<br>• Be aware that AI models may reflect and amplify biases present in their training data. |

### *Guidelines for Reviewers & Editors*

| Role | Specific Actionable Guidelines |
|---|---|
| Reviewers | • Shift focus to higher-order critiques: significance, novelty, methodological rigor, and argument strength, as AI can mask basic writing errors.<br>• Be vigilant for AI generation weaknesses: homogenized prose, lack of depth, factual inaccuracies, and generic phrasing.<br>• Do not use AI to analyse confidential manuscripts unless the journal's policy explicitly permits it. |
| Editors & Journals | • Develop clear, enforceable, and specific AI policies aligned with international standards (ICMJE, COPE).<br>• Implement a standardized AI disclosure section in the submission process.<br>• Educate reviewers and editorial staff on detecting and evaluating AI-assisted manuscripts.<br>• Focus on content quality and integrity rather than outright prohibition, while defining clear boundaries for unacceptable use. |

### *Guidelines for Institutions & Educators*

| Role | Specific Actionable Guidelines |
|---|---|
| Universities & Institutions | • Integrate AI literacy and ethics into curricula. Teach students how to use AI responsibly and critically.<br>• Develop clear institutional policies on AI use for research and assessment, promoting equitable access to tools.<br>• Redesign assessments to value process, critical reflection, and original thought (e.g., oral defences, in-class writing). |
| Educators | • Use AI-generated text as a teaching tool for comparative analysis and rhetorical awareness.<br>• Explicitly teach rhetorical strategies that AI often lacks, such as counter-argumentation and authorial presence.<br>• Embed the value of authorial voice and critical thinking into assignment rubrics and feedback. |

## Quick-Reference Checklist for Authors

Before submission, ask yourself:

- **AUTHORSHIP:** Have I ensured that AI is not listed as an author?
- **DISCLOSURE:** Have I written a clear disclosure statement for all non-routine AI uses, specifying the tool, purpose, and extent?
- **VERIFICATION:** Have I meticulously fact-checked, edited, and taken full responsibility for every part of the manuscript, including AI-generated content?
- **VOICE:** Does the final manuscript sound like *my* work? Have I preserved my analytical rigor and authorial voice?
- **INTEGRITY:** Am I confident that the work is original, accurate, and free from plagiarism?
- **PRIVACY:** Have I avoided inputting any sensitive or confidential data into AI tools?
- **POLICY:** Have I checked and adhered to the specific AI policy of my target journal?