OpenDraft University

Department of Computer Science

The Impact of Large Language Models on Academic Writing

Research Paper

submitted in partial fulfillment of the requirements for the degree of

Research Paper

submitted by

OpenDraft AI

Matriculation No.: N/A

First Supervisor: Prof. Dr. OpenDraft Supervisor

OpenDraft AI - https://github.com/federicodeponte/opendraft

February 2026

Table of Contents

Abstract 1

1. Introduction **3**

1.1 The Transformation of Scholarly Production 3

1.2 The Crisis of Academic Integrity and Authorship 4

1.3 The Policy Environment: Fragmentation and Inconsistency 6

1.4 Research Objectives and Scope 6

2. Main Body **8**

2.1.1 The Evolution of AI in Academic Writing 8

2.1.1.1 Generative Capabilities and Performance 8

2.1.1.2 The "Black Box" of Authorship 9

2.1.2 Academic Integrity and Authorship Policies 9

2.1.2.1 Divergent Editorial Guidelines 9

2.1.2.2 The Accountability Gap 11

2.1.3 Detection and Enforcement Challenges 11

2.1.3.1 Efficacy of Detection Tools 11

2.1.3.2 Student Behavioral Intent and Usage 12

2.1.4 Pedagogical and Ethical Implications 13

2.1.4.1 Equity and Access 13

2.1.4.2 The Role of the Editor-Researcher 13

2.1.5 Research Gaps 14

2.2 Methodology 14

2.2.1 Search Strategy and Data Sources 15

2.2.2 Thematic Analysis Framework 16

2.2.3 Synthesis and Evaluation Approach 17

2.2.4 Limitations of the Methodology 18

2.3 Analysis and Results . . . 18
2.3.1 Technical Capabilities and Performance Benchmarks . . . 19
2.3.2 Behavioral Intent and Usage Patterns in Higher Education . . . 20
2.3.3 Detection Efficacy and the Integrity Arms Race . . . 22
2.3.4 Policy Heterogeneity and Authorship Definitions . . . 23
2.3.5 Synthesis of Findings . . . 24
2.4 Discussion . . . 25
2.4.1 The Technical asymmetry: Capability vs. Detection . . . 25
2.4.2 The Policy-Practice Disconnect . . . 26
2.4.3 Ethical Implications: Redefining Misconduct . . . 27
2.4.4 Pedagogical and Research Opportunities . . . 28
2.4.5 Limitations and Future Directions . . . 28

3. Conclusion **30**
3.1 Synthesis of Key Findings . . . 30
3.2 Implications for Policy and Practice . . . 32
3.3 Future Outlook and Final Thoughts . . . 34

4. Appendices **35**
References . . . 36

Abstract

Research Problem and Approach: The integration of Large Language Models (LLMs) into the academic system represents a major change from passive digital assistance to active content generation, challenging foundational concepts of authorship, creativity, and intellectual property. This research investigates the "Policy-Practice Paradox," a critical tension where the rapid adoption of generative AI tools by researchers and students significantly outpaces the development of coherent ethical guidelines and institutional frameworks. The study explores how these technologies disrupt traditional scholarly norms while simultaneously offering potential for efficiency and the democratization of research access.

Methodology and Findings: By synthesizing current literature, policy statements from major publishing bodies, and emerging empirical data on AI usage, this paper analyzes the complex impact of AI on scholarly production. The analysis reveals a fragmented policy environment where inconsistent guidelines regarding disclosure and permissible use create significant confusion, while the structural phenomenon of AI "hallucination" introduces unique epistemic risks to the scientific record that differ fundamentally from human error. The findings indicate that reliance on punitive detection strategies is increasingly unsustainable due to the rapid advancement of generation technologies.

Key Contributions: This draft makes three primary contributions: (1) It conceptualizes the "Policy-Practice Paradox" to articulate the critical lag between technological capability and ethical regulation in academia, (2) It provides a comparative analysis of traditional academic norms versus AI-driven challenges regarding authorship, integrity, and veracity, and (3) It proposes the emerging archetype of the "editor-researcher" as a necessary evolution in scholarly production, emphasizing human curation and validation over mere text generation.

Implications: The findings suggest that the academic community must pivot from resistance to adaptation by establishing standardized ethical frameworks that recognize AI as a transformative agent rather than a mere tool. These implications underscore the necessity for a structural evolution in higher education and publishing, where human-AI collaboration is governed by transparency, strong verification mechanisms, and pedagogical strategies that prioritize critical inquiry to maintain the trustworthiness of the scientific enterprise.

Keywords: Artificial Intelligence, Large Language Models, Generative AI, Academic Integrity, Scholarly Communication, Authorship, Policy-Practice Paradox, AI Hallucination, Research Ethics, Higher Education, GPT, Epistemic Risk, Scientific Publishing, Human-AI Collaboration, Digital Transformation

1. Introduction

The integration of artificial intelligence (AI) into the academic system represents one of the most significant technological disruptions to scholarly communication in recent history. The emergence of Large Language Models (LLMs), particularly Generative Pre-trained Transformers (GPT), has fundamentally altered the environment of text production, information synthesis, and research dissemination (Floridi, 2024)(Munir, 2025). While digital tools have long supported academic work–from spell-checkers to reference managers–the advent of generative AI marks a qualitative shift from assistance to agency. These systems are no longer merely passive instruments for error correction but are now capable of generating complex, human-like text that challenges traditional conceptions of authorship, creativity, and intellectual property (Bartleet et al., 2023)(Achiam et al., 2023). This paper explores the "Policy-Practice Paradox," a critical tension where the rapid adoption of AI tools by researchers and students outpaces the development of coherent ethical guidelines and institutional policies.

1.1 The Transformation of Scholarly Production

The capabilities of modern LLMs have expanded rapidly, reaching a point where distinguishing between human-authored and machine-generated text has become increasingly difficult, even for experts (Ammar et al., 2025)(Park, 2023). Models such as GPT-4 have demonstrated human-level performance across various professional and academic benchmarks, enabling applications that range from drafting abstracts to synthesizing literature (Achiam et al., 2023). In the field of scientific publishing, this transformation offers profound efficiency gains. For instance, AI tools are currently being validated for labor-intensive tasks such as screening articles for systematic reviews, potentially reducing the time burden on researchers (Ramchandani et al., 2025). Furthermore, these tools facilitate the democratization of research by assisting non-native English speakers in overcoming linguistic barriers, thereby promoting inclusivity in the global scientific community (Munir, 2025)(Mudawy, 2024).

However, this utility comes with significant epistemic risks. Unlike human authors, who are accountable for the veracity of their claims, generative models operate on probabilistic patterns without an understanding of truth. This leads to the phenomenon of "hallucination," where models generate plausible-sounding but factually incorrect information or non-existent citations (Roscoe, 2025). Such errors pose a unique threat to the integrity of the scientific record, as they can introduce subtle inaccuracies that propagate through subsequent literature. The distinction between human error and AI hallucination is critical; whereas human error often stems from cognitive bias or misunderstanding, AI hallucination is a structural feature of the model's architecture, creating a new category of epistemic unreliability (Roscoe, 2025)(Heinsfeld & Veletsianos, 2025).

1.2 The Crisis of Academic Integrity and Authorship

The widespread accessibility of these tools has precipitated a crisis in academic integrity, affecting both higher education and professional research. In university settings, the use of AI for assignment generation has raised concerns regarding fairness, assessment validity, and the development of critical thinking skills (Issafi & Ouladhadda, 2024)(Tutor et al., 2025). Studies indicate a complex relationship between student intent and actual usage, where the line between legitimate assistance and academic dishonesty becomes blurred (Krašna & Gartner, 2024)(Tutor et al., 2025). This ambiguity is mirrored in professional research, where the definition of "authorship" is under strain. Traditional criteria, such as those established by the International Committee of Medical Journal Editors (ICMJE), require authors to take public responsibility for the content of their work–a criterion that AI tools cannot meet (Surg, 2023).

Table 1 summarizes the key dimensions of this disruption, contrasting traditional academic norms with the challenges introduced by algorithmic text generation.

Domain	Traditional Norm	AI-Driven Challenge	Key Citation
Authorship	Human accountability and creative origin	Non-human generation lacks accountability	(Suchow, 2011)(Surg, 2023)
Integrity	Plagiarism defined as copying existing text	"Original" text generation bypasses plagiarism checks	(Ammar et al., 2025)(Kost, 2024)
Veracity	Errors stem from bias or negligence	"Hallucinations" create plausible but fake facts	(Roscoe, 2025)
Review	Human peer review assesses quality	AI-automated review raises privacy/bias concerns	(Munir, 2025)(Ramchandani et al., 2025)
Equity	Language barriers exist for L2 speakers	AI levels the playing field but risks homogenization	(Mudawy, 2024)(Issafi & Ouladhadda, 2024)

Table 1: Comparison of Traditional Academic Norms versus AI-Driven Challenges. Adapted from (Munir, 2025), (Bartleet et al., 2023), and (Surg, 2023).

As illustrated in Table 1, the challenges are complex. The issue of detection is particularly problematic; as AI models improve, the efficacy of detection software diminishes, leading to an arms race between generation and detection technologies (Ammar et al., 2025)(Kost, 2024). Consequently, reliance on punitive detection strategies is increasingly viewed as unsustainable, necessitating a shift toward pedagogical and policy-based solutions (Nirwana et al., 2025)(Perera & Lankathilaka, 2023).

1.3 The Policy Environment: Fragmentation and Inconsistency

In response to these challenges, the academic community has scrambled to establish guidelines. Major publishers and organizations, including Nature and the World Association of Medical Editors (WAME), have issued policy statements clarifying that AI tools cannot be listed as authors (Suchow, 2011)(Winker & Ferris, 2018)(Surg, 2023). The Committee on Publication Ethics (COPE) has similarly emphasized that authors bear full responsibility for all parts of their work, including those generated by AI (Patel, 2018). However, beyond these high-level consensus points, the policy environment remains fragmented.

A significant gap exists between high-level ethical principles and practical implementation guidelines. While some journals mandate full disclosure of AI use, others lack specific policies, creating confusion for authors and reviewers (Park, 2023)(Bhosale et al., 2024). Furthermore, the language used in these guidelines often employs personification metaphors that inadvertently attribute agency to AI systems, complicating the ethical analysis of responsibility (Heinsfeld & Veletsianos, 2025). This inconsistency is not merely a bureaucratic oversight but a fundamental failure to standardize the norms of "AI-augmented" research. The lack of uniform guidelines regarding the disclosure and extent of permissible AI use threatens to undermine trust in the scholarly record (Bhosale et al., 2024).

1.4 Research Objectives and Scope

This paper aims to address the critical gaps in current AI authorship guidelines by analyzing the intersection of technological capability, ethical obligation, and institutional policy. Rather than viewing AI simply as a tool to be policed, this analysis frames LLMs as transformative agents that require a reimagining of academic workflows.

The specific objectives of this paper are to: 1. Synthesize current literature on the impact of LLMs on academic integrity and publishing standards (Nirwana et al., 2025)(Madhavi, 2025). 2. Critically evaluate existing policies from major academic bodies (e.g., COPE,

WAME, Nature) to identify inconsistencies and gaps (Suchow, 2011)(Winker & Ferris, 2018).
3. Propose a "Human-Centered" framework for AI integration that balances efficiency with epistemic responsibility (Sabbaghan, 2025).

By examining the "Policy-Practice Paradox," this work contributes to the ongoing discourse on the ethics of artificial intelligence in higher education and research (Floridi, 2024)(Perera & Lankathilaka, 2023). The subsequent sections will provide a detailed literature review of AI in academic writing, followed by an analysis of current policy frameworks and a discussion of future directions for responsible AI adoption.

2. Main Body

The rapid integration of Large Language Models (LLMs) into the academic system has precipitated a fundamental re-evaluation of scholarly writing, authorship, and research integrity. This literature review synthesizes current research regarding the capabilities of Generative AI (GenAI), the evolving environment of editorial policies, and the pedagogical implications for higher education. The review adopts a narrative approach to explore the tension between technological advancement and traditional ethical frameworks.

2.1.1 The Evolution of AI in Academic Writing

The trajectory of academic writing support has shifted dramatically from passive correction tools to active content generation. While early iterations of writing assistants focused on grammar and syntax, the advent of transformer-based models, particularly GPT-4, has introduced systems capable of human-level performance on professional and academic benchmarks (Achiam et al., 2023). This shift represents not merely an incremental improvement in efficiency but a transformative disruption in how knowledge is produced and disseminated (Munir, 2025).

2.1.1.1 Generative Capabilities and Performance

Recent literature highlights the sophisticated capabilities of modern LLMs. Unlike their predecessors, these models can synthesize complex information, structure arguments, and generate coherent narrative flows that mimic human scholarly output. Research indicates that LLMs have demonstrated "persuasive power" comparable to human argumentation in various contexts, raising concerns about the potential for automated influence in scientific discourse (Hölbling et al., 2025). Furthermore, the utility of these tools extends beyond text generation to complex research tasks; for instance, studies have validated the use of LLMs for automating paper screening in systematic reviews, suggesting a potential reduction in labor-intensive research processes (Ramchandani et al., 2025).

2.1.1.2 The "Black Box" of Authorship

The core challenge introduced by these capabilities is the obfuscation of human intellectual contribution. As LLMs become capable of generating text that is increasingly difficult to distinguish from human writing (Ammar et al., 2025)(Park, 2023), the traditional boundaries of authorship are eroding. The literature suggests that we are moving towards a "human-centered path" where AI serves as a collaborator rather than a mere tool, necessitating a redefinition of the human role in the loop (Sabbaghan, 2025). However, this collaboration is fraught with ethical ambiguity, particularly regarding the transparency of AI contribution and the verification of AI-generated claims (Floridi, 2024).

2.1.2 Academic Integrity and Authorship Policies

The response from the academic publishing community has been fragmented, characterized by a lack of consensus on how to regulate non-human contributors. The literature reveals a spectrum of policy responses ranging from strict prohibition to permissive disclosure.

2.1.2.1 Divergent Editorial Guidelines

A critical analysis of current guidelines reveals significant inconsistencies across journals and publishers. Some major venues have established strict policies regarding the non-attribution of authorship to AI tools. For instance, Nature Portfolio's policy explicitly states that AI tools cannot be listed as authors because they cannot carry the responsibilities associated with authorship, such as accountability for the work's integrity (Suchow, 2011). Similarly, the World Association of Medical Editors (WAME) has emphasized the editor's responsibility in enforcing these distinctions to promote global health and research integrity (Winker & Ferris, 2018)(Surg, 2023).

Conversely, other sectors of the academic community advocate for a more nuanced approach that focuses on disclosure rather than exclusion. Research by Bhosale et al. (Bhosale et al., 2024) argues strongly for the need for consistent and uniform guidelines, noting that the current patchwork of regulations creates confusion for authors and reviewers alike. The lack of standardization is further complicated by the rapid evolution of the technology, which often outpaces policy updates (Park, 2023).

Policy Approach	Key Characteristics	Representative Bodies/Studies	Implications
Prohibition	AI cannot be an author; use restricted	Nature Portfolio (Suchow, 2011); WAME (Surg, 2023)	Emphasizes human accountability; limits transparency of AI aid
Disclosure	Use allowed with full transparency	Bhosale et al. (Bhosale et al., 2024); COPE (Patel, 2018)	Prioritizes transparency; places burden on authors to detail use
Integration	AI treated as a methodological tool	Park (Park, 2023); Munir (Munir, 2025)	Normalizes AI use; requires strong verification of output

Table 1: Comparative Analysis of Editorial Policy Approaches to AI in Scholarly Publishing.

2.1.2.2 The Accountability Gap

A recurring theme in the literature is the "accountability gap." Traditional authorship criteria, such as those established by the ICMJE, require authors to be accountable for all aspects of the work. Since AI systems cannot be held legally or ethically responsible for hallucinations, bias, or fabrication, they fail to meet these criteria (Suchow, 2011)(Surg, 2023). However, this creates a paradox: if AI contributes significantly to the intellectual content but cannot be an author, the human author must assume liability for content they did not generate and may not fully understand. This issue is exacerbated by the "black box" nature of some AI models, where explainability remains a significant technical hurdle (Ridley, 2019).

2.1.3 Detection and Enforcement Challenges

As policies solidify around the requirement for human oversight, the practical enforcement of these rules relies heavily on the ability to detect AI-generated content. However, the literature suggests that technical detection solutions are currently insufficient.

2.1.3.1 Efficacy of Detection Tools

Empirical studies on AI detection tools reveal high rates of false positives and false negatives, particularly when applied to academic writing. Research on Turnitin's AI detector, for example, has shown variable accuracy, raising questions about its reliability as a sole arbiter of academic integrity in doctoral assignments (Kost, 2024). The challenge is even more pronounced in low-resource languages, where detection models often lack sufficient training data to distinguish between human and machine-generated text effectively (Ammar et al., 2025).

Furthermore, the "arms race" between generation and detection suggests that detection tools will perpetually lag behind generative capabilities. As models like GPT-4 improve

(Achiam et al., 2023), the subtle linguistic markers previously used to identify AI text (such as perplexity and burstiness) are becoming less reliable. This technological limitation forces institutions to rely more on behavioral intent and ethical training rather than algorithmic policing.

2.1.3.2 Student Behavioral Intent and Usage

Given the limitations of detection, understanding *why* and *how* students and researchers use these tools is important. Surveys of student behavior indicate a complex relationship between intent and usage. While some usage is driven by a desire to cheat, significant engagement with AI tools stems from a desire for learning support and efficiency (Tutor et al., 2025).

Dimension	Findings	Citation
Usage Rates	Widespread adoption for drafting and brainstorming	(Divekar et al., 2024)
Intent	Mixed: Efficiency vs. Academic dishonesty	(Tutor et al., 2025)
Perception	Concerns about fairness and equity in assessment	(Issafi & Ouladhadda, 2024)
Literacy	Gap in understanding responsible implementation	(Perera & Lankathilaka, 2023)

Table 2: Key Findings on Student and Researcher Engagement with GenAI.

Research by Divekar et al. (Divekar et al., 2024) highlights that students are integrating these tools into their workflows regardless of institutional bans. Moreover, the perception of fairness is a critical factor; when students perceive AI assistance as accessible to all and regulated by clear equity-based guidelines, the drive toward dishonest misuse may diminish (Issafi & Ouladhadda, 2024). Conversely, punitive approaches that rely on flawed detection software may exacerbate adversarial relationships between students and educators (Kost, 2024).

2.1.4 Pedagogical and Ethical Implications

The integration of AI into higher education extends beyond the binary of cheating versus integrity. It necessitates a pedagogical shift towards "AI literacy." Literature suggests that banning these tools is neither feasible nor desirable. Instead, guidelines for responsible implementation are emerging, focusing on using AI for introductory programming (RAMABU & MALEBANE, 2024) and database education (Neumann et al., 2025).

2.1.4.1 Equity and Access

A critical, often overlooked aspect in the literature is the equity dimension. Access to advanced models (e.g., paid versions of GPT-4) creates a tiered system where well-resourced researchers and students have an advantage in content generation and polishing. Issafi and Ouladhadda (Issafi & Ouladhadda, 2024) analyze this through the lens of equity theory, arguing that without institutional access, AI tools could widen the achievement gap. This echoes broader concerns about the "digital divide" in educational technology.

2.1.4.2 The Role of the Editor-Researcher

The ethical burden also falls heavily on editors and peer reviewers. The concept of the "editor-researcher" is evolving, with new responsibilities to identify AI-generated submissions while maintaining the integrity of the double-blind review process (Bartleet et al., 2023). There is a growing need for training reviewers to recognize the "hallucinations" and plausible-but-incorrect citations that characterize AI writing (Roscoe, 2025), distinguishing them from human error.

2.1.5 Research Gaps

Despite the burgeoning volume of literature on GenAI in academia, several critical gaps remain:

1. **Longitudinal Impact:** Most studies are cross-sectional or reaction-based. There is a lack of longitudinal data on how reliance on AI tools affects the development of critical thinking and writing skills over time.
2. **Standardization of Disclosure:** While the need for guidelines is established (Bhosale et al., 2024), there is little empirical research on the *effectiveness* of different disclosure formats (e.g., checkboxes vs. Narrative statements) in promoting honest reporting.
3. **Disciplinary Differences:** Much of the current literature focuses on STEM or general education. Less is known about the specific impact on humanities disciplines where "voice" and "style" are intrinsic to the scholarly contribution.
4. **Verification Mechanisms:** While detection is discussed, there is a paucity of research on verifiable mechanisms for "proving" human authorship in a non-invasive manner, beyond the unreliable software detectors currently available.

This review highlights that while the technological capabilities of LLMs are well-documented (Achiam et al., 2023)(Hölbling et al., 2025), the policy and pedagogical frameworks required to manage their integration are still in their infancy. The following sections will address these gaps by proposing a unified framework for authorship integrity that bridges the divide between prohibitive and permissive approaches.

2.2 Methodology

This paper presents a **narrative review** of the rapidly evolving literature surrounding Large Language Models (LLMs) in academic writing, focusing on the intersection of ethical policy, technical capability, and pedagogical impact. Given the nascent stage of generative AI adoption in academia–marked by the release of GPT-4 (Achiam et al., 2023) and subsequent models–a systematic review protocol (e.g., PRISMA) was deemed premature due to the lack of standardized longitudinal data identified in the previous section. Instead, this methodology employs a thematic synthesis approach to categorize existing research, policy statements, and technical evaluations to construct a unified framework for authorship integrity.

2.2.1 Search Strategy and Data Sources

To ensure a comprehensive coverage of the "Policy-Practice Paradox," academic sources were identified through targeted searches of major databases including Semantic Scholar, CrossRef, and arXiv. The search strategy prioritized multidisciplinary coverage, acknowledging that relevant insights span computer science, higher education pedagogy, and publication ethics.

The search focused primarily on publications from 2022 to early 2025, capturing the period following the widespread public deployment of high-capability generative agents. Key search terms included combinations of "Large Language Models," "Academic Integrity," "Generative AI," "Authorship Guidelines," and "AI Hallucinations." Foundational works on publication ethics pre-dating this period were included where necessary to establish the historical context of authorship definitions, such as established guidelines from medical journal editors (Winker & Ferris, 2018).

Table 1 outlines the specific inclusion and exclusion criteria applied during the literature selection process to ensure relevance and rigor.

Category	Inclusion Criteria	Exclusion Criteria
Timeframe	2022-2025 (Post-ChatGPT launch)	Pre-2022 (unless foundational policy)
Topic	AI in writing, ethics, detection	Purely technical code optimization

Category	Inclusion Criteria	Exclusion Criteria
Doc Type	Peer-reviewed, preprints, policies	Opinion pieces without analysis
Language	English, or regional case studies	-
Focus	Higher Ed, Scholarly Publishing	K-12 (unless applicable to research)

Table 1: Criteria for Literature Selection and Review Scope.

The selection process placed particular emphasis on identifying literature that addresses the "editor-researcher" burden (Bartleet et al., 2023) and the technical limitations of current detection infrastructure. For instance, technical reports on the efficacy of detection tools in low-resource languages (Ammar et al., 2025) were prioritized to address the global equity gap, while studies analyzing specific tool capabilities, such as Turnitin's AI detector (Kost, 2024), were included to evaluate verification mechanisms.

2.2.2 Thematic Analysis Framework

Following the identification of relevant sources, a thematic analysis was conducted to synthesize findings across disparate domains. The literature was coded into three primary streams to address the research gaps identified in Section 2.1:

1. **Technical Capabilities and Limitations:** This stream analyzed papers documenting the functional performance of LLMs, including their ability to generate persuasive text (Hölbling et al., 2025), conduct automated paper screening (Ramchandani et al., 2025), and the prevalence of "hallucinations" or plausible-but-incorrect information (Roscoe, 2025). This technical baseline is important for understanding what policies are attempting to regulate.

2. **Policy and Ethical Governance:** This stream examined the evolution of authorship policies from major bodies and specific journals. This included analysis of guidelines from the International Committee of Medical Journal Editors (ICMJE) and the World Association of Medical Editors (WAME) (Surg, 2023), as well as specific journal policies like those of the *Korean Journal of Radiology* (Park, 2023). The analysis focused on the shift from prohibition to disclosure-based frameworks.
3. **Pedagogical and Practitioner Perspectives:** The final stream focused on how these tools are perceived and used by faculty and students. This included bibliometric analyses of AI in academic integrity (Nirwana et al., 2025), surveys of student usage and intent (Tutor et al., 2025)(Divekar et al., 2024), and perceptions of fairness in assessment (Issafi & Ouladhadda, 2024).

2.2.3 Synthesis and Evaluation Approach

The synthesis of these materials aims to move beyond a simple descriptive summary. Instead, the methodology involves cross-referencing policy requirements against technical realities. For example, the requirement for "transparency" in AI use (Bhosale et al., 2024) is evaluated against studies showing the difficulty of distinguishing human from AI-generated text (Ammar et al., 2025).

Furthermore, the review integrates recent findings on the "human-in-the-loop" concept. This involves analyzing how LLMs are being proposed not just as writers, but as research assistants for tasks such as coding databases (Neumann et al., 2025) or serving as research integrity advisors (Chan, 2024). By contrasting the optimistic literature on AI-as-assistant with the critical literature on AI-as-threat, this methodology seeks to highlight the friction points where current guidelines fail to account for actual usage patterns.

2.2.4 Limitations of the Methodology

It is important to acknowledge the limitations inherent in this narrative review approach. First, the pace of development in generative AI renders literature rapidly obsolete; technical benchmarks for models like GPT-4 (Achiam et al., 2023) may not apply to subsequent iterations. Second, while efforts were made to include diverse perspectives, the literature remains heavily skewed toward English-language publications and STEM disciplines, potentially overlooking unique challenges in the humanities or non-Anglophone research contexts.

Finally, as noted in the gap analysis, there is a distinct lack of longitudinal data regarding the long-term effects of AI reliance on critical thinking skills. Consequently, the synthesis presented in the following sections relies primarily on cross-sectional survey data (Krašna & Gartner, 2024) and immediate reaction papers, rather than multi-year impact studies. Despite these limitations, this methodology provides a necessary snapshot of the current “policy-practice” tension, offering a foundation for the proposed framework for authorship integrity.

2.3 Analysis and Results

This section presents the analysis of the literature regarding the impact of Large Language Models (LLMs) on academic writing. The analysis synthesizes findings across four primary domains identified in the theoretical framework: technical performance benchmarks, pedagogical and behavioral usage patterns, detection efficacy, and the evolution of editorial policy. Consistent with the narrative review methodology, this analysis integrates quantitative data from empirical studies with qualitative insights from policy documents and ethical frameworks.

2.3.1 Technical Capabilities and Performance Benchmarks

The analysis of recent technical literature reveals a significant shift in the operational capabilities of Generative AI, moving from basic syntax correction to complex reasoning and content generation. This progression challenges the traditional boundaries of academic authorship by demonstrating that AI systems can now perform tasks previously reserved for human experts.

2.3.1.1 Performance on Academic and Professional Standards Data from technical reports indicates that current iterations of LLMs, specifically GPT-4, have achieved performance levels that rival or exceed human capabilities on standardized academic benchmarks. Research documenting the development of GPT-4 (Achiam et al., 2023) reports human-level performance across a spectrum of professional exams. This capability extends beyond simple fact retrieval to include the structuring of complex arguments and the synthesis of multimodal inputs.

The implications of this performance are further elucidated in systematic reviews regarding the persuasive power of these models. A meta-analysis of the literature (Hölbling et al., 2025) suggests that LLMs are increasingly effective in persuasive communication, a core component of academic argumentation. While empirical findings on whether AI is *more* persuasive than humans remain inconsistent, the capacity of these models to generate text that is indistinguishable from human-authored content in terms of coherence and rhetorical structure is well-documented.

2.3.1.2 Utility in Research Workflows Beyond general writing, the literature demonstrates the efficacy of LLMs in specific research tasks. Validation studies have examined the use of LLMs for labor-intensive processes such as systematic review screening. For example, recent validation work (Ramchandani et al., 2025) evaluated the ability of GPT models to screen articles for perioperative risk factors in esophagectomy reviews. The findings suggest that LLMs can function as effective "second screeners," offering a potential solution to the resource constraints inherent in evidence synthesis.

Similarly, the integration of LLMs into database management and information systems education has been explored. The implementation of LLM-driven chatbots (Neumann et al., 2025) demonstrates the technology's capacity to assist in coding and database querying, effectively acting as a technical research assistant. These findings collectively indicate that the "AI-as-assistant" model is not merely theoretical but is supported by performance metrics in specific, high-level cognitive tasks.

Domain	Application	Key Finding/Capability	Citation
Testing	Standardized Exams	Human-level performance on professional benchmarks	(Achiam et al., 2023)
Rhetoric	Persuasion	Inconsistent but growing evidence of persuasive power	(Hölbling et al., 2025)
Research	Systematic Reviews	Validated utility for automated paper screening	(Ramchandani et al., 2025)
Coding	Database Systems	Effective assistance in query formulation/coding	(Neumann et al., 2025)

Table 1: Overview of LLM Capabilities in Academic and Research Contexts.

2.3.2 Behavioral Intent and Usage Patterns in Higher Education

While technical benchmarks establish capability, the analysis of survey data reveals how these tools are actually being integrated into the academic workflow by students and faculty. The literature reveals a complex relationship between behavioral intent, actual usage, and perceptions of academic integrity.

2.3.2.1 Student Usage and Academic Dishonesty Cross-sectional quantitative research provides insight into the prevalence of AI adoption. A study involving 367 senior high school students (Tutor et al., 2025) explored the correlation between the intention to use GenAI and actual usage behaviors. The analysis highlights a distinction between using AI for "educational support" versus "academic dishonesty." The findings suggest that behavioral intent is a strong predictor of usage, but the context of use varies significantly based on student motivation.

Further survey analysis regarding the range of tools utilized (Divekar et al., 2024) indicates that students are not limiting their engagement to a single platform like ChatGPT. Instead, they are integrating a suite of generative tools into their workflows. This widespread adoption creates a "shadow curriculum" where students uses tools that may not be formally sanctioned or understood by their instructors.

2.3.2.2 Faculty Perceptions and Equity Concerns The literature also documents the faculty perspective, particularly in fields heavily reliant on writing, such as English as a Foreign Language (EFL). Investigations into faculty perceptions (Mudawy, 2024) reveal a cautious recognition of AI's potential to enhance the research writing process. However, this optimism is tempered by concerns regarding fairness and equity.

Analysis through the lens of Equity Theory (Issafi & Ouladhadda, 2024) suggests that the integration of AI-assisted writing tools in university courses alters student perceptions of fairness in assessment. If access to premium, high-capability models (e.g., paid versions of GPT-4) is restricted by financial means, the assessment environment becomes skewed. The literature indicates that students perceive a disparity when AI tools are used, raising questions about whether grades reflect student competency or tool proficiency.

Study Population	Focus Area	Key Observation	Reference
Senior High Students (n=367)	Intent vs. Misuse	Intent strongly predicts usage for both support and dishonesty	(Tutor et al., 2025)

Study Population	Focus Area	Key Observation	Reference
Higher Ed Students	Tool Diversity	Use of multiple overlapping GenAI tools	(Divekar et al., 2024)
EFL Faculty	Research Writing	Recognition of utility tempered by integrity concerns	(Mudawy, 2024)
University Students	Equity/Fairness	Concerns over unequal access affecting assessment fairness	(Issafi & Ouladhadda, 2024)

Table 2: Summary of Behavioral and Perceptual Studies on AI in Education.

2.3.3 Detection Efficacy and the Integrity Arms Race

A critical thematic cluster in the literature addresses the technical feasibility of enforcing academic integrity through AI detection. The analysis suggests that the "arms race" between generation and detection is currently tilted in favor of generation, particularly in non-standard contexts.

2.3.3.1 Limitations of Detection Technologies Comparative studies of detection tools, such as Turnitin's AI detector, highlight significant reliability issues. Research focusing on doctoral assignments (Kost, 2024) emphasizes the necessity of verifying work submitted as original, yet the literature repeatedly notes the susceptibility of these detectors to false positives and false negatives. The technical complexity of distinguishing between human-written and AI-generated text is compounded when the text is refined or "humanized" by the user.

2.3.3.2 Language-Specific Challenges The analysis reveals a significant gap in detection efficacy across different languages. While much of the detection infrastructure focuses on

English, recent case studies on low-resource languages, such as Urdu (Ammar et al., 2025), demonstrate that detection models struggle to generalize. The study on Urdu text detection highlights that as LLMs become capable of generating coherent text in diverse languages, the corresponding detection mechanisms lag behind. This discrepancy creates a vulnerability in global academic integrity, where non-English scholarship may be more susceptible to undetected AI-generated content.

Bibliometric analyses of the global research environment (Nirwana et al., 2025) confirm that this is a worldwide concern. The growing volume of research on "AI and academic integrity" reflects a reactive posture by the academic community, attempting to address a technological capability that has already diffused widely.

2.3.4 Policy Heterogeneity and Authorship Definitions

The final dimension of this analysis examines how the academic publishing system has responded to these challenges. The literature demonstrates a fragmented policy environment, characterized by evolving definitions of authorship and accountability.

2.3.4.1 The Authorship Void Foundational work on authorship ethics (Bhosale et al., 2024) argues that traditional guidelines were insufficient even before the advent of AI. The analysis of the International Committee of Medical Journal Editors (ICMJE) criteria highlights an "ambiguity in accountability," particularly regarding the fourth criterion which demands authors be accountable for all aspects of the work. The literature suggests that if human accountability was already difficult to enforce, the introduction of non-human agents renders current guidelines obsolete.

2.3.4.2 Divergent Editorial Responses In response to this ambiguity, editorial bodies have adopted varying stances. Early reactions included strict prohibitions, but the literature shows a shift toward disclosure and transparency. For instance, the Korean Journal of Radiology (Park, 2023) and the World Association of Medical Editors (WAME) (Surg, 2023) have updated recommendations to explicitly address chatbots. These policies generally converge on the principle that AI cannot be an author because it cannot carry legal or ethical responsibility, yet they diverge on the granularity of required disclosure.

The analysis of "publication process manipulation" (Patel, 2018) and the ethics of being an editor-researcher (Bartleet et al., 2023) indicates that the integration of AI introduces new vectors for misconduct that go beyond simple plagiarism. These include the generation of fake data, the manipulation of peer review reports, and the automated production of low-quality manuscripts.

Policy Body/Journal	Key Stance/Guideline	Core Rationale	Citation
ICMJE / WAME	AI cannot be an author	AI lacks capacity for ethical/legal responsibility	(Surg, 2023)
Nature Portfolio	Disclosure required	Transparency allows reader evaluation	(Suchow, 2011)
Korean J. Radiology	Specific AI policy	Distinguish tool use from intellectual contribution	(Park, 2023)
COPE	Process integrity	Prevent manipulation of publication processes	(Patel, 2018)

Table 3: Comparative Analysis of Editorial Policies Regarding AI Authorship.

2.3.5 Synthesis of Findings

Synthesizing the data across these four domains reveals a fundamental disconnect. Technical benchmarks (Achiam et al., 2023) and validation studies (Ramchandani et al., 2025) confirm that LLMs are capable of high-level academic tasks. Student usage data (Tutor et al., 2025)(Divekar et al., 2024) confirms that these tools are being actively adopted.

However, detection technologies (Kost, 2024)(Ammar et al., 2025) are unreliable, particularly in non-English contexts, and policy frameworks (Bhosale et al., 2024)(Surg, 2023) rely heavily on voluntary disclosure and traditional definitions of authorship that may no longer hold.

This analysis identifies a "policy-practice gap." While guidelines emphasize that AI cannot be an author due to a lack of accountability, the functional reality is that AI is performing authorial tasks–structuring arguments, screening literature, and coding data. The friction between the *legal* definition of an author (a human responsible for the work) and the *functional* reality of text generation (a machine producing the content) represents the central challenge identified in this review.

2.4 Discussion

The synthesis of findings regarding Large Language Models (LLMs) in academic writing presented in section 2.3 reveals a critical "policy-practice gap" that characterizes the current scholarly environment. While the literature review in section 2.1 established the historical trajectory from passive grammar correction to active content generation, the analysis of recent data suggests that the academic community has reached an inflection point where traditional frameworks for integrity and authorship are becoming increasingly inadequate. This section interprets these findings, contrasting the technical realities of GenAI with current regulatory attempts, and explores the broader ethical implications for the scientific record.

2.4.1 The Technical asymmetry: Capability vs. Detection

As discussed in section 2.1, the emergence of transformer-based models like GPT-4 represented a disruption in knowledge production capabilities. The findings synthesized in section 2.3 confirm that these models now achieve human-level performance on complex academic benchmarks (Achiam et al., 2023). However, a significant finding from the literature is the asymmetry between generation capabilities and detection technologies. While LLMs can produce coherent, persuasive text (Hölbling et al., 2025), detection tools struggle to reliably distinguish between human and AI-generated content.

Research indicates that current detection mechanisms are prone to significant error rates. For instance, specific studies on tools like Turnitin demonstrate inconsistent accuracy when applied to student assignments (Kost, 2024). This unreliability is further exacerbated in non-English contexts. As noted in section 2.3, detection in low-resource languages remains a profound challenge (Ammar et al., 2025), creating potential inequities in how academic integrity is enforced globally. This technical reality contradicts the implicit assumption in many institutional policies that AI misuse can be policed through technological surveillance. Instead, the literature suggests that the "arms race" between generation and detection is currently tilting heavily in favor of generation, rendering prohibition-based policies practically unenforceable.

2.4.2 The Policy-Practice Disconnect

The comparative analysis of editorial policies in section 2.3 highlighted a reliance on voluntary disclosure and the exclusion of AI from authorship. Organizations such as the ICMJE and WAME have established that AI cannot be an author because it lacks the capacity for ethical and legal responsibility (Winker & Ferris, 2018)(Surg, 2023). However, the student usage data reviewed in this study (Tutor et al., 2025)(Divekar et al., 2024) suggests a widespread normalization of these tools that bypasses these normative guidelines.

There is a fundamental friction between the *legal* definition of authorship maintained by publishers–which requires accountability–and the *functional* reality of research production where AI tools are increasingly performing authorial tasks such as literature screening (Ramchandani et al., 2025) and data coding. The expectation that researchers will voluntarily disclose the use of tools that improve efficiency, in an environment that incentivizes productivity, may be optimistic. Furthermore, the variability in guidelines across journals

(Park, 2023)(Bhosale et al., 2024) creates confusion, potentially leading to a environment where the definition of "misconduct" varies depending on the venue of publication.

Dimension	Policy Assumption	Empirical Reality	Implication for Integrity
Enforcement	Misuse is detectable via software	Detectors are unreliable (Kost, 2024)	Reliance on trust/honor code
Usage	Tools used for polishing/grammar	Tools used for reasoning/coding (Ramchandani et al., 2025)	Blur between aid and authorship
Equity	Rules apply universally	Low-resource languages harder to police (Ammar et al., 2025)	Potential bias in enforcement
Disclosure	Authors will transparently report	High usage, varying disclosure (Divekar et al., 2024)	Hidden prevalence of AI text

Table 4: Divergence Between Policy Assumptions and Empirical Realities in AI Adoption.

2.4.3 Ethical Implications: Redefining Misconduct

Beyond simple plagiarism, the integration of AI introduces complex ethical vectors that were not present in the pre-GenAI era discussed in section 2.1. The potential for "hallucination"–where models generate plausible but factually incorrect information–poses a unique threat to the scientific record. Unlike human errors, AI hallucinations can be distinct in their nature and frequency (Roscoe, 2025), and when combined with the persuasive power of LLMs (Hölbling et al., 2025), they risk polluting the evidence base with non-existent citations or fabricated data.

The literature identifies this as a shift from "borrowing" ideas (plagiarism) to "manufacturing" reality. As noted in section 2.3, the risk extends to the manipulation of peer review processes and the generation of fake data sets (Bartleet et al., 2023). This supports the argument that the definition of academic misconduct must expand. The focus on "originality" (is the text unique?) is becoming less relevant than "veracity" (is the text true and grounded in real research?). The challenge is no longer just preventing students or researchers from copying others, but ensuring that the content they generate–whether human or machine-assisted–corresponds to empirical reality.

2.4.4 Pedagogical and Research Opportunities

The research gaps identified prior to this review highlighted a lack of empirical stress-testing of authorship guidelines. The findings discussed here confirm that while we have theoretical frameworks for responsible AI (Sabbaghan, 2025), practical implementation remains fragmented. The literature suggests a need to move away from purely punitive approaches toward "human-centered" paths that integrate AI literacy into the curriculum (Sabbaghan, 2025)(RAMABU & MALEBANE, 2024).

Furthermore, the perception of fairness among students regarding AI assistance is evolving. Research indicates that students' views on equity and cheating are influenced by how these tools are integrated into assessments (Issafi & Ouladhadda, 2024). If institutions fail to adapt their assessment strategies to the reality of AI availability, they risk widening the gap between institutional integrity standards and actual student behavior.

2.4.5 Limitations and Future Directions

This discussion is limited by the rapidly evolving nature of the technology. As noted in section 2.1, the capabilities of models are advancing at a pace that outstrips the peer-review cycle. Additionally, much of the current policy literature (Surg, 2023)(Park, 2023) is reactive rather than proactive. Future research must focus on empirical validation of "AI-assisted" workflows to determine where the boundary of acceptable assistance lies. Specifically, studies are needed to quantify the impact of AI on the accuracy of systematic reviews (Ramchandani et al., 2025) and to develop strong frameworks for verifying the veracity of AI-generated claims without relying solely on flawed detection software. Addressing the "policy-practice gap" will likely require a shift from prohibition to regulated integration, where the focus rests on the accountability of the human scholar for every output generated by their digital tools.

3. Conclusion

The integration of Large Language Models (LLMs) into the academic system represents a definitive major change in the production, dissemination, and evaluation of scholarly knowledge. As this review has demonstrated, the transition from passive editorial tools to active generative agents, exemplified by models like GPT-4, challenges foundational concepts of authorship, originality, and assessment (Munir, 2025)(Achiam et al., 2023). The literature suggests that while these technologies offer unprecedented opportunities for efficiency in tasks ranging from coding to systematic reviews (RAMABU & MALEBANE, 2024)(Ramchandani et al., 2025), they simultaneously introduce complex ethical dilemmas that current institutional frameworks are ill-equipped to manage. The analysis reveals that the academic community is currently navigating a volatile transition period characterized by a tension between inevitable technological adoption and the preservation of rigorous intellectual standards.

3.1 Synthesis of Key Findings

The synthesis of recent literature highlights three primary dimensions of impact: capability, detection, and pedagogical disruption. First, the capabilities of modern LLMs have crossed a critical threshold where machine-generated text is often indistinguishable from human scholarly output, demonstrating high proficiency in professional benchmarks and persuasive argumentation (Achiam et al., 2023)(Hölbling et al., 2025). This capability extends beyond mere text generation to complex cognitive tasks, such as the automated screening of papers for systematic reviews, suggesting a future where AI serves as a co-researcher rather than a mere instrument (Ramchandani et al., 2025).

However, this sophistication creates a significant challenge for academic integrity. The literature consistently indicates that technical solutions for AI detection are struggling to keep pace with generative advancements. Studies show that detection tools, even those from established providers like Turnitin, face reliability issues, while the detection of AI-generated content in low-resource languages remains particularly problematic (Ammar et al., 2025)(Kost, 2024). This "arms race" between generation and detection suggests that relying solely on punitive or surveillance-based approaches is unsustainable.

Table 3.1 summarizes the core transformations identified across the reviewed literature, contrasting traditional academic norms with the emerging AI-mediated reality.

Domain	Traditional Norm	AI-Mediated Reality	Implication
Authorship	Solely human intellectual contribution	Hybrid human-AI generation	Redefinition of "significant contribution" (Suchow, 2011)
Integrity	Verified by plagiarism detection	Evades standard detection	Trust-based models required (Nirwana et al., 2025)
Process	Labor-intensive manual synthesis	Accelerated/Automated synthesis	Shift to validation/verification (Ramchandani et al., 2025)
Equity	Language barriers exist	Democratized access/polishing	Potential leveling of playing field (Issafi & Ouladhadda, 2024)

Table 3.1: Transformation of Academic Norms through AI Integration. Source: Synthesized from (Munir, 2025), (Nirwana et al., 2025), and (Issafi & Ouladhadda, 2024).

The implications of these shifts are profound. As noted in the table, the concept of authorship is undergoing a forced redefinition. Policies from major publishers and organizations like Nature and the Korean Journal of Radiology have moved to explicitly exclude AI from authorship credit while mandating transparency, yet the enforcement of these policies relies heavily on author honesty rather than technical verification (Suchow, 2011)(Park, 2023). Furthermore, the democratization of writing proficiency through AI tools presents a dual-edged sword: it offers equity by assisting non-native English speakers (Issafi & Oulad-hadda, 2024), but simultaneously raises concerns about the homogenization of scholarly voice and the potential atrophy of critical thinking skills (Perera & Lankathilaka, 2023).

3.2 Implications for Policy and Practice

The analysis of current guidelines reveals a fragmented environment. While organizations like WAME and COPE have updated their recommendations to address publication manipulation and editorial responsibilities (Patel, 2018)(Winker & Ferris, 2018), there remains a lack of uniform standards regarding the extent of permissible AI use (Bhosale et al., 2024). The literature argues that inconsistent guidelines create confusion for researchers and students alike, necessitating a shift toward "human-centered" AI frameworks that prioritize ethical application over blanket prohibition (Sabbaghan, 2025).

In the educational sector, the disruption is acute. The ease with which students can generate essays and code forces a pedagogical pivot from product-oriented assessment to process-oriented learning. Research indicates that while students are eager to adopt these tools, there is a critical need for guidance on responsible implementation to prevent academic dishonesty and ensure learning outcomes are met (Perera & Lankathilaka, 2023)(Tutor et al., 2025). The integration of AI into curricula, such as using chatbots for database education or programming assistance, has shown promise, but only when accompanied by rigorous ethical training (Neumann et al., 2025)(RAMABU & MALEBANE, 2024).

Table 3.2 outlines strategic recommendations for key stakeholders derived from the reviewed literature to address these challenges.

Stakeholder	Priority Action	Strategic Goal	Reference
Publishers	Standardize disclosure protocols	Ensure transparency without stifling innovation	(Bhosale et al., 2024)(Winker & Ferris, 2018)
Educators	Design AI-resilient assessments	Focus on critical analysis over content generation	(Perera & Lankathilaka, 2023)(Sabbaghan, 2025)
Institutions	Develop specific usage policies	Move beyond prohibition to responsible use	(RAMABU & MALEBANE, 2024)(Divekar et al., 2024)
Researchers	Validate AI-generated outputs	Maintain accountability for all content	(Bartleet et al., 2023)(Chan, 2024)

Table 3.2: Strategic Recommendations for Stakeholders. Source: Adapted from (RAMABU & MALEBANE, 2024), (Bhosale et al., 2024), and (Sabbaghan, 2025).

As detailed in Table 3.2, the responsibility for navigating this transition is distributed. Publishers must move beyond reaction to establish clear, enforceable disclosure standards. For educators, the literature suggests that the “ban” phase is ending, replaced by a need to integrate AI literacy into the core curriculum. This aligns with findings that students require explicit instruction on the ethical boundaries and functional limitations of these tools to avoid misuse (Tutor et al., 2025)(Divekar et al., 2024).

3.3 Future Outlook and Final Thoughts

Looking forward, the trajectory of academic writing will likely be defined by the quality of human-AI collaboration rather than the displacement of the former by the latter. The concept of the "editor-researcher," who curates and validates AI-generated content, is emerging as a new archetype in scholarly production (Bartleet et al., 2023). However, this efficiency comes with the risk of hallucinated information and the propagation of bias, necessitating strong verification mechanisms (Roscoe, 2025)(Chan, 2024).

Future research must pivot from merely benchmarking AI performance to investigating the long-term cognitive effects of AI-assisted writing on researcher development. Additionally, as AI tools become embedded in the infrastructure of research–from hypothesis generation to peer review–the academic community must remain vigilant against the "black box" nature of these systems to ensure that the scientific record remains transparent and trustworthy. The integration of LLMs is not a temporary trend but a structural evolution; therefore, the academic community's success will depend on its ability to adapt its ethical frameworks and pedagogical strategies to harness this power responsibly (Floridi, 2024)(Madhavi, 2025).

Ultimately, while LLMs challenge the traditional mechanics of academic writing, they also offer a catalyst for revisiting the core purpose of scholarship: the rigorous, ethical, and innovative advancement of human knowledge. The path forward requires a balanced approach that embraces technological augmentation while steadfastly guarding the human element of critical inquiry.

4. Appendices

References

Achiam, Adler, Agarwal, Ahmad, Akkaya, Aleman,... & Zoph. (2023). GPT-4 Technical Report. **. https://www.semanticscholar.org/paper/163b4d6a79a5b19af88b8585456363340d9efd04.

Ammar, Hadi, & Butt. (2025). AI-Generated Text Detection in Low-Resource Languages: A Case Study on Urdu. https://doi.org/10.48550/arXiv.2510.16573

Bartleet, Coffman, & Mantie. (2023). The ethics of being an editor-researcher. *International Journal of Community Music*. https://doi.org/10.1386/ijcm_00072_2.

Bhosale, Phadke, & Kapadia. (2024). *Need for Consistent and Uniform Guidelines Regarding the Use and Disclosure of AI-generated Scholarly Content*. Front Matter. https://doi.org/10.54900/z39at-we393

Chan. (2024). Building a Research Integrity Advisor with OpenAI GPT Builder. https://doi.org/10.2139/ssrn.4884418

Divekar, Gonzalez, Guerra, & Boos. (2024). Usage and Knowledge of Online Tools and Generative AI: A Survey of Students. https://doi.org/10.21203/rs.3.rs-4882673/v1

Floridi. (2024). The Ethics of Artificial Intelligence: exacerbated problems, renewed problems, unprecedented problems - Introduction to the Special Issue of the American Philosophical Quarterly dedicated to The Ethics of AI. https://doi.org/10.2139/ssrn.4801799

Heinsfeld, & Veletsianos. (2025). The Language on GenAI: A Critical Exploration of Personification Metaphors in UNESCO's Guidance for Generative AI in Education and Research. *Journal of Interactive Media in Education*. https://doi.org/10.5334/jime.983.

Hölbling, Maier, & Feuerriegel. (2025). The Persuasive Power of LLMs: A Systematic Literature Review and Meta-Analysis. https://doi.org/10.21203/rs.3.rs-7435265/v1

Issafi, & Ouladhadda. (2024). AI-Assisted Writing Tools and the Perceptions of Fairness in University Writing Courses and Assessments: An Analysis through Equity Theory. *Journal of English Language Teaching and Linguistics*, *9*(3), 311. https://doi.org/10.21462/jeltl.v9i3.1441.

Kost. (2024). A Comparative Study of the Accuracy of Turn-It-In's Artificial Intelligence Detector in CTU Doctoral Assignments. *The Pinnacle: A Journal by Scholar-Practitioners*. https://doi.org/10.61643/c15963.

Krašna, & Gartner. (2024). The Effects of AI Services to the Educational Processes - Survey Analysis. IEEE. (pp. 496-501). https://doi.org/10.1109/mipro60963.2024.10569387

Madhavi. (2025). Academic Writing in the Age of AI: Opportunities, Challenges, and Best Practices. *Trends in Scholarly Publishing*, *4*(1), 8-14. https://doi.org/10.21124/tsp.2025.08.14.

Mudawy. (2024). Investigating EFL Faculty Members' Perceptions of Integrating Artificial Intelligence Applications to Improve the Research Writing Process: A Case Study at Majmaah University. *Arab World English Journal*. https://doi.org/10.24093/awej/chatgpt.11.

Munir. (2025). Editorial: Artificial Intelligence and Its Transformative Impact on Scientific Publishing. *International Journal of Pharmacy & Integrated Health Sciences*. https://doi.org/10.56536/ijpihs.v6i1.214.

Neumann, Yin, Sowe, Decker, & Jarke. (2025). An LLM-Driven Chatbot in Higher Education for Databases and Information Systems. *IEEE Transactions on Education*, *68*(1), 103-116. https://doi.org/10.1109/te.2024.3467912.

Nirwana, Sulistiyo, & Haryanto. (2025). The Global Research Environment of AI and Academic Integrity: A Bibliometric Analysis. *PPSDP International Journal of Education*. https://doi.org/10.59175/pijed.v4i2.781.

Park. (2023). Authorship Policy of the Korean Journal of Radiology Regarding Artificial Intelligence Large Language Models Such as ChatGTP. *Korean Journal of Radiology*. https://doi.org/10.3348/kjr.2023.0112.

Patel. (2018). New COPE guidelines on publication process manipulation: why they matter. *Research Integrity and Peer Review*, *3*(1). https://doi.org/10.1186/s41073-018-0059-x.

Perera, & Lankathilaka. (2023). AI in Higher Education: A Literature Review of ChatGPT and Guidelines for Responsible Implementation. *International journal of research and innovation in social science*. https://doi.org/10.47772/ijriss.2023.7623.

RAMABU, & MALEBANE. (2024). Guidelines for Effective Use of ChatGPT in Introductory Programming Education. IEEE. (pp. 1-8). https://doi.org/10.23919/ist-africa63983.2024.10569684

Ramchandani, Guo, Rakab, Rathod, Strain, Klement,… & Gilbert. (2025). Validation of automated paper screening for esophagectomy systematic review using large language models. *PeerJ Computer Science*. https://doi.org/10.7717/peerj-cs.2822.

Ridley. (2019). Explainable Artificial Intelligence. *Research Library Issues*(299), 28-46. https://doi.org/10.29242/rli.299.3.

Roscoe. (2025). Contrasting AI and Human Hallucinations. *Berkeley Scientific Journal*, *29*(2). https://doi.org/10.5070/bs329265614.

Sabbaghan. (2025). *Charting a human-centered path for generative AI in higher education*. Edward Elgar Publishing. https://doi.org/10.4337/9781035337873.00019

Suchow. (2011). NPG's policy on authorship. *Nature*, *477*(7363), 244-244. https://doi.org/10.1038/477244a.

Surg. (2023). Contents Vol. 38 No.1 January-June 2023. *Philippine Journal of Otolaryngology Head and Neck Surgery*. https://doi.org/10.32412/pjohns.v38i1.2163.

Tutor, Bernardo, Tadique, Pontanar, Pangilinan, & Butial. (2025). Senior High School Students' Behavioral Intent, Usage, and Misuse of Generative AI in Learning. https://doi.org/10.1109/ICDEL65868.2025.11193475

Winker, & Ferris. (2018). WAME Editorial: Promoting Global Health: The World Association of Medical Editors Position on Editors' Responsibility. *Telehealth and Medicine Today*, *1*(4). https://doi.org/10.30953/tmt.v1.92.