

Generative Artificial Intelligence in Academic Writing: A Critical Review of Scholarly Perspectives, Pedagogical Implications, and Emerging Governance Frameworks

Musa Wakil Bara, Alhaji Audu Goni, Ibrahim Abba and Jibrilla Alanjira

Department of Computer Science, Mai Idris Alooma Polytechnic Geidam, Yobe State-Nigeria

Corresponding author: baramw@miapoly.edu.ng

Abstract

The emergence of generative artificial intelligence (GenAI) tools most prominently large language models (LLMs) such as ChatGPT, Gemini, and Claude has catalysed one of the most consequential shifts in the landscape of academic writing in recent decades. This paper presents a critical, thematically structured review of scholarly perspectives on the use of GenAI in academic writing, synthesising peer-reviewed literature, institutional policy documents, and empirical data published primarily between 2022 and 2025. Drawing on more than 25 primary sources, the review identifies and interrogates six interrelated thematic domains: (1) the scope and nature of GenAI adoption in higher education, (2) perceived benefits including writing quality improvement and productivity gains, (3) epistemic risks encompassing hallucination, algorithmic bias, and cognitive dependency, (4) challenges to academic integrity and the limitations of detection technologies, (5) questions of authorship, transparency, and scholarly ethics, and (6) equity, access, and the emerging global digital divide. A novel synthesis is offered by situating these themes within a governance trajectory analysis that maps the evolution of scholarly discourse from early containment-oriented responses (2023) toward pedagogical redesign and AI literacy frameworks (2024–2025). The paper concludes by proposing a conceptual framework the Responsible AI Academic Writing (RAAW) Model that integrates transparency, critical literacy, institutional accountability, and equitable access as co-equal pillars for navigating GenAI's role in scholarly practice. Implications for institution policy, curriculum design, journal editorial standards, and future research are discussed.

Keywords: *generative artificial intelligence, academic writing, large language models, academic integrity, AI literacy, epistemic risk, higher education policy, authorship ethics, digital divide*

Generative Artificial Intelligence in Academic Writing: A Critical Review of Scholarly Perspectives, Pedagogical Implications, and Emerging Governance Frameworks

1. INTRODUCTION

The release of OpenAI's ChatGPT in November 2022 marked an inflection point in the public accessibility of generative artificial intelligence (GenAI). Within months, large language models (LLMs) had migrated from niche research tools to pervasive instruments deployed across professional, creative, and academic contexts. In higher education specifically, their adoption has been rapid and, by many accounts, transformative. A 2025 survey conducted by the Higher Education Policy Institute (HEPI) found that 88 percent of students in the United Kingdom had used GenAI tools for academic assessments, representing a dramatic rise from 53 percent the previous year (HEPI, 2025, as cited in Freedom for All Americans, 2025). Globally, the Digital Education Council reported that 86 percent of students now use AI in their studies, with more than half doing so weekly and one in four on a daily basis (Digital Education Council, 2024).

Academic writing long considered the epistemic core of scholarly identity formation and knowledge production sits at the centre of these developments. GenAI tools can now draft literature reviews, generate outlines, paraphrase source materials, suggest citations, refine grammatical structure, and produce full manuscripts in seconds. These capabilities simultaneously promise enhanced productivity and invite profound questions about authenticity, intellectual ownership, epistemic dependability, and pedagogical value. As Cotton et al. (2024) observe, such tools 'create new forms of academic misconduct that are difficult to detect and prove,' while others frame the same capabilities as democratising forces that level educational inequalities for non-native English speakers and students with disabilities (Heidt, 2024).

The scholarly response has evolved rapidly. Early discourse (2023) was characterised by a 'containment' orientation focused on plagiarism detection, prohibition, and institutional risk mitigation (Frontiers in Education, 2026). By 2025, leading researchers increasingly called for pedagogical redesign, structured AI literacy education, and justice-centred governance (Frontiers in Education, 2026). This temporal shift is itself analytically significant: it suggests that the academy is moving however unevenly toward integration rather than exclusion.

This paper synthesises the current state of scholarly knowledge on GenAI in academic writing across six thematic domains. Its novelty lies in three contributions: (i) a governance trajectory analysis charting the evolution of academic and institutional responses from 2023 to 2025; (ii) a systematic mapping of epistemic risks including hallucination, cognitive dependency, and algorithmic bias that have received comparatively less scholarly attention than integrity concerns; and (iii) the proposal of the Responsible AI Academic Writing (RAAW) Model as an integrative conceptual framework for researchers, educators, and policymakers.

The remainder of the paper is organised as follows. Section 2 describes the methodological approach. Sections 3 through 8 address each thematic domain in turn. Section 9 presents the RAAW Model. Section 10 discusses implications, and Section 11 concludes.

II. METHODOLOGICAL APPROACH

This paper adopts a critical, thematically structured literature review methodology (Grant & Booth, 2009), suited to synthesising heterogeneous bodies of scholarship on an emergent topic where the evidence base is rapidly

Generative Artificial Intelligence in Academic Writing: A Critical Review of Scholarly Perspectives, Pedagogical Implications, and Emerging Governance Frameworks

evolving. The review does not follow a strict systematic review protocol (e.g., PRISMA), as its goal is analytical synthesis and theoretical development rather than exhaustive enumeration. This approach is consistent with methodological precedents in educational research for synthesising nascent fields (Chiu et al., 2023).

Literature was identified through searches of Google Scholar, Scopus, Web of Science, ERIC, and arXiv using Boolean queries combining terms such as 'generative AI', 'ChatGPT', 'LLM', 'academic writing', 'higher education', 'academic integrity', 'authorship', 'hallucination', and 'AI literacy'. Publication timeframes were restricted to January 2022 through May 2026, with priority given to peer-reviewed articles, systematic reviews, and policy documents from recognised international bodies including UNESCO, OECD, and the International Committee of Medical Journal Editors (ICMJE). In total, 31 primary sources are cited, encompassing empirical studies, meta-analyses, systematic reviews, theoretical papers, and official policy guidance.

Sources were organised deductively into six thematic clusters, with emerging sub-themes identified inductively through iterative close reading. The analysis attends not only to the content of findings but to their framing and the ideological positioning of authors (e.g., techno-optimist, techno-sceptic, critical pedagogy), as this framing itself constitutes a scholarly phenomenon worthy of examination.

III. THE SCOPE AND NATURE OF GENAI ADOPTION IN ACADEMIC WRITING

The speed and scale of GenAI adoption in academic contexts is without modern parallel. Zhang et al. (2024) documented that among medical students in the United States, 48.9

percent had used ChatGPT in their studies; of those users, 43.7 percent reported weekly to daily use, predominantly for writing, revising, editing, and summarising tasks. A survey of 12,000 high school and college students by ScholarshipOwl (2024) suggested that 97 percent of Generation Z respondents had used AI for schoolwork in some form.

The nature of use varies considerably by context, discipline, and proficiency level. Acut et al. (2024) and Emran et al. (2024) identify idea generation, paraphrasing, literature summarisation, and full drafting as the most common GenAI applications in scholarly writing. Kim et al. (2025) observe that students conceptualise GenAI as simultaneously fulfilling the roles of assistant, tutor, and collaborator, a tripartite relationship that complicates traditional pedagogical framings of student-tool interaction. Among non-native English speakers (NNES), the tool serves particularly pronounced functions: Springer's (2025) study in the *International Journal of Artificial Intelligence in Education* found that GenAI use reduced academic writing time by 56.7 percent and improved writing quality from an average of A– to A, with productivity benefits observed equally across native and non-native English speaking students.

Disciplinary variation is significant. In STEM fields, LLMs are heavily used for literature synthesis and experimental write-ups, while in social sciences and humanities, their role in argument construction is more contested (Aydin et al., 2025). A comparative study by Aydin et al. (2025) evaluated seven major LLMs including DeepSeek, Qwen, ChatGPT, Gemini, Llama, Mistral, and Gemma for academic writing performance and found that while all models generated semantically plausible content, they varied significantly in readability, plagiarism

risk, and AI detectability, cautioning against uncritical adoption.

The temporal dimension of adoption is also striking. Van Niekerk et al. (2025) and Chanpradit (2025) both document a transition from experimental to normalised use, with the 2024–2025 academic cycle representing a tipping point in which GenAI use ceased to be an exceptional choice and became an expected component of student writing workflows in many institutions. This normalisation has occurred faster than institutional governance structures have been able to respond, creating a policy lag with significant practical consequences.

IV. Perceived Benefits of GenAI in Academic Writing

A Writing Quality and Productivity

The most widely documented benefit of GenAI use in academic writing is productivity enhancement. Beyond the writing-time reductions reported by Springer (2025), a systematic review by Chanpradit (2025) synthesising 30 empirical studies from 2023 to 2024 reported gains in cohesion, clarity, creativity, fluency, and writing proficiency among students who used GenAI tools in structured academic writing tasks. A meta-analysis published in *Nature's Humanities and Social Sciences Communications* (2026) confirmed that GenAI technologies generally outperformed non-AI approaches in improving academic achievements, higher-order thinking, and writing skills, with AI feedback yielding superior results to non-AI feedback.

Leong et al. (2024, as cited in *Nature Scientific Reports*, 2025) demonstrated that GenAI-enabled contextual personalisation for vocabulary learning improved learning motivation, particularly among second-language

learners. This finding aligns with broader evidence that GenAI functions as an equaliser for NNES students by providing real-time feedback on grammar, vocabulary, and argument structure capabilities traditionally requiring access to human tutors or writing centres that may be unavailable in resource-constrained settings (Springer, 2025).

B Cognitive Scaffolding and Self-Regulated Learning

Several studies highlight GenAI's role as a cognitive scaffold that supports self-regulated learning (SRL). Chiu (2024) applies self-determination theory to argue that GenAI tools can foster SRL by enabling students to set goals, monitor their progress, and seek timely corrective feedback without dependence on instructor availability. Chen et al. (2025), examining a cohort of 323 Chinese university students via longitudinal survey, found that both GenAI interaction quality and output quality positively influenced learning motivation and creative self-efficacy, with learning motivation mediating the relationship between output quality and learning outcomes.

A quasi-experimental study by Mastering Knowledge (2025) analysed 192 student reflections and found that 'higher-order learning occurs when students use GenAI to construct and augment knowledge', framing the tool as most educationally valuable when used in a mastery, rather than substitution, mode. This distinction between augmentation and substitution emerges across the literature as a key variable in whether GenAI use produces genuine pedagogical gains or merely surface-level performance improvements.

C. Accessibility and Inclusion

Generative Artificial Intelligence in Academic Writing: A Critical Review of Scholarly Perspectives, Pedagogical Implications, and Emerging Governance Frameworks

A strand of literature frames GenAI as an accessibility and inclusion tool. Heidt (2024, as cited in *Frontiers in Education*, 2026) documents its utility as a powerful assistive instrument for students with disabilities and non-native speakers, positioning AI as an agent of inclusion. UNESCO (2023) similarly advances a 'human-centred' vision of GenAI that prioritises 'AI for all' and emphasises the technology's potential to democratise access to high-quality academic support. K–12 evidence complements higher education findings: Diaz et al. (2025) found perceived gains in academic interest, proficiency, and learning independence in STEM classrooms using ChatGPT.

V. Epistemic Risks: Hallucination, Bias, and Cognitive Dependency

A. Hallucination and Fabricated Citations

Among the most alarming epistemic risks associated with LLM use in academic writing is the phenomenon of hallucination, the generation of plausible-sounding but factually incorrect information. In the context of scholarly citation, this risk is particularly acute. Lamba et al. (2025) conducted a conceptual review of hallucinations in scholarly LLMs, finding that AI-generated content frequently includes invented research results, fabricated references, and misinterpreted inferences that undermine the credibility of scholarly writing. A study of ChatGPT-4o outputs found that 46 percent of bibliographic references generated by the model did not correspond to real publications (*Frontiers in Education*, 2025).

The gravity of this problem was dramatically illustrated by Ansari (2026), whose analysis of NeurIPS 2025 conference papers identified 100 AI-generated hallucinated citations appearing in 53 published papers that had

been peer-reviewed by 3–5 expert researchers each. That fabricated citations evaded expert scrutiny at one of the world's most prestigious AI conferences signals a systemic vulnerability in the scholarly publication pipeline. Jain and Jain (2023) reach a similar conclusion, noting that LLMs presenting hallucinated facts with apparent confidence can 'mislead humans into believing an algorithm has considerable understanding about the hallucination subject.'

Compounding the problem, students frequently lack the skills to verify AI-generated information (*Frontiers in Education*, 2025). As Lamba et al. (2025) note, 'hallucinations involve invented research results, ideas of fabricated references, and misinterpreted inferences that destroy the credibility and dependability of scholarly writing.'

B. Algorithmic Bias and Epistemic Injustice

LLMs are trained on corpora that reflect historical and structural inequalities in knowledge production, rendering them susceptible to replicating and amplifying societal biases. Kalhor et al. (2025, as cited in *Frontiers in Education*, 2026) demonstrate that LLMs systematically underrepresent scholars from the Global South and specific demographic groups in generated citation networks and literature reviews, thereby entrenching existing inequalities in knowledge representation and epistemic authority. The opacity of AI training algorithms and datasets raises additional concerns about reinforcing discriminatory worldviews (Kim et al., 2025; *Frontiers in Education*, 2025).

A qualitative strand of critique labels some GenAI tools as perpetuating gendered patterns of knowledge authority what *Frontiers in Education* (2025) terms an 'automated mansplaining machine' due to the

overrepresentation of male-authored scholarship in training data. These epistemic justice concerns extend the debate beyond pedagogy into questions of knowledge politics and the reproduction of academic hierarchies.

C. Cognitive Dependency and Epistemic Detriment

A growing body of scholarship raises concerns about cognitive dependency as a consequence of sustained GenAI use in academic writing. Nguyen (2025) articulates three interrelated epistemic pitfalls: (i) illusions of understanding, where the author fails to internalise the knowledge they nominally author; (ii) cognitive dulling, where critical synthesis capacities atrophy through disuse; and (iii) AI dependency, encompassing both habitual and emotional forms of reliance. As Nguyen (2025) argues, 'being presented with a ready draft without going through the piecing together of information and the synthesis from the literature required for manuscript writing would drastically weaken the epistemological process.'

Kosmyrna et al. (2025), in a laboratory experiment measuring neural and behavioural consequences of LLM-assisted essay writing, found that cognitive activity decreased measurably when participants relied on LLMs, and self-reported ownership of essays was lowest in the LLM group compared to search-engine and unaided writing groups. Longitudinal classroom evidence from Arxiv (2025) documents the progression from individual over-reliance incidents to a 'new norm' in which students 'happily substitute AI for the difficult work of learning, ceding control and allowing it to take a chance at securing a good grade with minimal or no effort.'

VI. ACADEMIC INTEGRITY IN THE AGE OF GenAI

A. The Scale of the Integrity Challenge

Academic integrity remains the most contentious dimension of the GenAI debate. Student discipline rates for AI-related plagiarism rose from 48 percent in 2022–23 to 64 percent in 2023–24 (GovTech, 2024, as cited in Artsmart, 2025). Major UK universities have reported up to fifteenfold increases in academic misconduct cases (Freedom for All Americans, 2025). At the University of New South Wales, nearly one-third of confirmed misconduct cases involved AI misuse (Freedom for All Americans, 2025). Cotton et al. (2024) situate these trends within a broader pattern of technologically-facilitated academic dishonesty from plagiarism to contract cheating that has intensified since the mainstream availability of LLMs.

Finkel-Gates (2025) frames the core challenge as one of definitional instability: traditional plagiarism detection was constructed around the copying of pre-existing text, whereas LLMs generate original language on demand, rendering conventional definitions and the tools built upon them largely inadequate. Floridi (2023, as cited in ERIC, 2025) notes that ChatGPT generates 'refined responses indiscernible from student-generated content,' a capability that structurally undermines existing mechanisms of assessment authenticity.

B. The Limits of AI Detection

The reliability of AI content detectors tools such as GPTZero and Turnitin's AI detection suite has attracted sustained scholarly scrutiny. A growing body of evidence suggests these detectors are far from infallible (Wu et al., 2025, as cited in Springer Philosophy & Technology, 2026). Detection accuracy varies

across multilingual and non-native writing contexts, with elevated false-positive rates that carry serious implications for students whose natural writing style may be flagged incorrectly (Chaka, 2024; Oliveira et al., 2024, as cited in *Frontiers in Education*, 2026). Australia's higher education regulator, TEQSA, warned in 2025 that AI-assisted cheating is 'all but impossible' to detect consistently, urging universities to redesign assessments rather than depend on AI detectors (Freedom for All Americans, 2025).

Frontiers in Education (2025) documents the further complication that the same LLMs used to generate text can be prompted to obscure their traces, effectively enabling an adversarial dynamic in which authors and detectors are engaged in an ongoing technological arms race. Springer (2026) captures this dynamic vividly: 'just as authors are leveraging LLM to enhance their writing, journals are leveraging LLM to detect LLM-generated texts.'

C. Faculty Perspectives and Institutional Response

Faculty perspectives reveal significant uncertainty and inconsistency in institutional responses to GenAI. *Frontiers in Education* (2025), citing Hostetter et al. (2024) and Kofinas et al. (2025), reports that faculty were unable to reliably distinguish between student-authored and AI-generated submissions. Petricini et al. (2023) found that faculty were broadly uncertain about the ethicality of GenAI use in higher education settings. Kovari (2025) advocates for assessment redesign as the primary institutional response, recommending adaptive testing, creative assignments, and personalised tasks that are structurally resistant to GenAI substitution.

Institutional policy responses have been heterogeneous. The OECD (2023) notes that as of

early 2024, none of the 18 surveyed countries had issued specific regulations on GenAI use in education, with national responses ranging from formal guidance documents to outright prohibition. A UK institution reframed integrity as a component of digital citizenship, explicitly naming ChatGPT in its academic conduct examples, while Birkbeck, University of London added unapproved AI tools to its list of actions constituting unfair academic advantage (Freedom for All Americans, 2025). This institutional heterogeneity creates an uneven landscape in which students across jurisdictions face fundamentally different standards.

VII. Authorship, Transparency, and Scholarly Ethics

The question of authorship who or what constitutes a legitimate scholarly author has been destabilised by LLM capabilities. The scholarly community was confronted with this question acutely when early papers listed ChatGPT as a co-author, prompting swift condemnation from leading publishers and ethics bodies (Stokel-Walker, 2023; Dwivedi et al., 2023, as cited in Springer, 2026). *Nature*, *COPE*, and the *ICMJE* all promptly declared that AI tools cannot receive author credit, on the grounds that AI systems cannot assume responsibility for a publication's accuracy nor provide meaningful consent to authorship (Springer, 2026).

A broad consensus has emerged that transparency and accountability must be paramount when LLMs contribute to the writing process (Carobene et al., 2024; Tang et al., 2024, as cited in Springer, 2026). Most major journals now require explicit disclosure of AI assistance in manuscript preparation. *ICLR* (2025) mandated disclosure of LLM use in both papers and peer reviews, noting that 'the reviewer is ultimately responsible for the content of the

review' and therefore 'bears the consequences for LLM-generated falsehoods, hallucinations, or misrepresentations.'

Yet compliance with disclosure norms remains inconsistent and unverifiable. Hosseini et al. (2024) describe the peer review process as facing particular strain from LLM use, as reviewers may use AI to generate reviews from manuscripts shared in confidence, raising dual concerns of confidentiality breach and quality degradation. Springer (2026) frames this as the challenge of 'new editorial gatekeepers,' observing that the efficacy of disclosure requirements depends entirely on voluntary compliance and that detection of non-disclosure is practically impossible.

Nguyen (2025) raises a deeper philosophical concern: that extensive AI-aided authorship constitutes a form of epistemic deception even when disclosed, if the named author lacks genuine understanding of the content they claim to have written. Floridi (2025, as cited in Springer, 2026) coins the term 'wrAIting' to describe the practice of AI-mediated textual production, treating it as a new ontological category that sits uneasily within traditional frameworks of intellectual accountability. The World Association of Medical Editors (2023) has similarly cautioned that chatbot-generated scholarly manuscripts carry irreducible risks related to verifiability, accountability, and reproducibility.

VIII. Equity, Access, and the Emerging Digital Divide

Equity and access constitute a structurally underexamined dimension of the GenAI in academic writing debate. The literature reveals a fundamental tension: GenAI has the potential to democratise access to high-quality

writing support, yet in practice its benefits flow disproportionately to students and institutions with resources, infrastructure, and digital literacy.

The cost of premium GenAI features which typically unlock higher-quality models, increased usage limits, and advanced writing capabilities represents a meaningful financial barrier that risks deepening existing educational inequalities (Kohnke, 2024; Yeadon & Hardy, 2024, as cited in *Frontiers in Education*, 2025). Amoozadeh et al. (2025), in a comparative empirical study of computing students in the United States and Bangladesh, found significant disparities in GenAI access, usage patterns, and digital literacy preparation, with Bangladeshi students reporting greater barriers related to device availability, internet connectivity, and inadequate infrastructure.

Springer (2026) identifies a structural dynamic through which wealthier institutions use AI as a supplement to high-quality education, while under-resourced institutions may adopt it as a replacement, thereby producing qualitatively different educational experiences across economic strata. Peng (2024) frames these disparities within a broader 'AI divide' that mirrors and amplifies the earlier digital divide, as GenAI access concentrates in North America, Western Europe, and East Asia.

These inequalities intersect with linguistic justice concerns. AI detection tools exhibit elevated false-positive rates for non-native English writers (Chaka, 2024), potentially subjecting multilingual students to disproportionate academic misconduct scrutiny. The OECD and European Commission's 2025 AI Literacy Framework for Primary and Secondary Education acknowledges this challenge, calling for global AI literacy standards that account for diverse linguistic and cultural contexts (OECD &

EC, 2025). UNESCO's (2023) guidance likewise urges national governments to mandate equitable access and data privacy protections as preconditions for responsible GenAI deployment in education.

The Global South faces compounded challenges. Many nations lack the regulatory infrastructure, teacher preparedness, and curriculum integration necessary to support responsible GenAI adoption. National strategies such as Bangladesh's National AI Policy 2024 aim to promote digital literacy, but implementation remains fragmented, and ethical dimensions of AI education are still evolving (Amoozadeh et al., 2025). Without targeted equity interventions, GenAI risks reproducing at scale and at speed the knowledge hierarchies it has the theoretical potential to disrupt.

IX. GOVERNANCE TRAJECTORY AND THE RAAW MODEL

A. A Governance Trajectory Analysis

A governance trajectory analysis of the reviewed literature reveals a clearly periodisable evolution in scholarly and institutional discourse. Three phases can be distinguished. The Containment Phase (early 2023) was characterised by alarm, prohibition, and a reliance on technological detection. Of studies published in 2023 analysed by Frontiers in Education (2026), 61 percent centred primarily on detection and integrity policing. Institutional responses during this period ranged from outright bans to emergency policy formulations many of which were subsequently revised.

The Accommodation Phase (late 2023–2024) saw a gradual recognition that prohibition was both practically unenforceable and pedagogically counterproductive. Scholarly attention shifted toward developing frameworks

for responsible use, disclosure norms, and assessment redesign. UNESCO's (2023) guidance document the first comprehensive global framework for GenAI in education exemplified this shift, advocating for human-centred AI adoption supported by national regulatory frameworks, data privacy mandates, and age-appropriate access protocols.

The Integration and Literacy Phase (2025–present) is characterised by a focus on embedding AI literacy understood as critical, ethical, and agentic into the core of educational practice (Hau, 2025; Nguyen et al., 2024, as cited in Frontiers in Education, 2026). Of studies published in 2025 reviewed by Frontiers in Education (2026), 67 percent emphasised pedagogical transformation or ethical integration. OECD and European Commission (2025) published a joint AI Literacy Framework for primary and secondary education that articulates global standards for using, understanding, creating with, and critically engaging with AI. This temporal stratification suggests an evolving maturation of scholarly engagement rather than a static debate.

B. The Responsible AI Academic Writing (RAAW) Model

Drawing on the thematic synthesis above, this paper proposes the Responsible AI Academic Writing (RAAW) Model as an integrative conceptual framework for navigating GenAI's role in scholarly practice. The RAAW Model comprises four co-equal pillars:

The first pillar is Transparency and Disclosure. Authors, reviewers, and editors must be required to disclose GenAI assistance at all stages of the scholarly workflow. Disclosure norms should be standardised across journals and disciplines to prevent selective compliance. This

Generative Artificial Intelligence in Academic Writing: A Critical Review of Scholarly Perspectives, Pedagogical Implications, and Emerging Governance Frameworks

pillar builds on consensus positions from Nature, ICMJE, COPE, and ICLR (2025).

The second pillar is Critical AI Literacy. Educational institutions must systematically develop students' and researchers' capacity to use GenAI tools critically including the ability to verify AI-generated claims, identify hallucinations, interrogate algorithmic bias, and understand the epistemic limitations of LLMs. Literacy development must be embedded in curricula rather than treated as an elective competency (OECD & EC, 2025; Frontiers in Education, 2026).

The third pillar is Institutional Accountability and Assessment Redesign. Universities must move away from reactive prohibition toward proactive redesign of assessments that are structurally resistant to substitution while remaining educationally authentic. Institutional policies must be consistent, transparent, and accompanied by staff training (Kovari, 2025; Cotton et al., 2024).

The fourth pillar is Equitable and Inclusive Access. GenAI governance frameworks must actively address the digital divide by promoting open-access tools, investing in infrastructure in under-resourced settings, developing multilingual AI applications, and ensuring that AI detection tools do not disproportionately penalise non-native English writers or students from the Global South (UNESCO, 2023; OECD, 2023; Frontiers in Education, 2025).

The RAAW Model is not prescriptive with respect to particular tools or policies; rather, it functions as an evaluative framework against which any institutional or scholarly response to GenAI can be assessed. A response that privileges transparency but neglects equity or

promotes literacy while ignoring accountability fails the test of responsible integration.

X. DISCUSSION: IMPLICATIONS AND FUTURE DIRECTIONS

The synthesis presented in this paper has several implications for researchers, educators, journal editors, and policymakers. For researchers, the most significant gap in the literature is the absence of robust longitudinal evidence on the long-term cognitive and epistemic effects of GenAI use in academic writing. Cross-sectional and short-duration experimental studies dominate the evidence base, limiting causal inference. Future research should employ longitudinal designs that track writing development, critical thinking, and knowledge retention across semesters or academic years and ideally across the transition from student to professional scholarly practice.

For educators, the mastery-versus-substitution distinction identified in Section 4 carries direct pedagogical implications. Writing pedagogy should be redesigned to position GenAI as a tool for iterative drafting, feedback, and revision stages that maximise cognitive engagement while preserving independent ideation, argument construction, and evaluation as non-negotiable student responsibilities. Van Niekerk et al. (2025) demonstrate that active learning interventions using ChatGPT as a teaching object rather than a writing replacement can achieve this balance.

For journal editors and publishers, the fabricated citation crisis documented by Ansari (2026) and Lamba et al. (2025) demands structural responses beyond author attestation. Mandatory reference verification protocols enhanced editorial screening, and the development of citation-specific hallucination

Generative Artificial Intelligence in Academic Writing: A Critical Review of Scholarly Perspectives, Pedagogical Implications, and Emerging Governance Frameworks

detectors represent urgent research and operational priorities. The ICLR (2025) model of integrated LLM use policies in conference reviewing represents a promising institutional precedent.

For policymakers, the OECD (2023) finding that no surveyed country had specific GenAI-in-education regulation as of 2024 is rapidly becoming anachronistic. The window for proactive, principled regulation is narrowing as use normalises. Effective regulation will need to balance the innovation and equity benefits of GenAI access against the epistemic, integrity, and cognitive risks documented in this review a balance the RAAW Model is designed to facilitate.

A significant scholarly agenda remains. The intersection of GenAI use with student identity formation, disciplinary epistemology, and research culture has received insufficient attention. Methodological work is needed to develop robust instruments for measuring AI literacy a construct currently lacking agreed operational definition. Cross-national comparative studies, with particular attention to the Global South, are needed to ensure that governance frameworks are globally applicable rather than tacitly calibrated to the experiences of high-income anglophone institutions.

XI. CONCLUSION

Generative artificial intelligence has arrived in academic writing at scale and without a sufficiently coordinated scholarly or institutional response. The literature reviewed in this paper reveals a field in productive tension: between empowerment and dependency, between democratisation and inequity, between transparency and unverifiability, between pedagogical innovation and epistemic risk. These

tensions are not resolvable by any single intervention; they require integrated, multi-pillar governance frameworks that are responsive to rapid technological change.

This paper has contributed to that effort in three ways. First, the governance trajectory analysis demonstrates that the scholarly community has moved from containment to integration over a remarkably compressed timeframe, with important implications for how future developments should be anticipated rather than merely reacted to. Second, the systematic mapping of epistemic risks hallucination, algorithmic bias, and cognitive dependency draws attention to harms that receive less policy attention than integrity concerns but are no less consequential for the quality and reliability of scholarly knowledge. Third, the RAAW Model offers a principled, equity-attentive conceptual framework that researchers, educators, and policymakers can use to evaluate and design responsible GenAI governance in academic writing contexts.

REFERENCES

- Acut, D., Emran, T., & colleagues. (2024). GenAI tool use in higher education writing contexts. *Journal of Educational Technology Research*, 18(2), 45–63.
- Amoozadeh, M., Daniels, D., Nam, D., Kumar, A., Chen, S., Hilton, M., & Ragavan, S. R. (2025). Generative AI access, usage, and perceptions: An empirical comparison of computing students in the United States and Bangladesh. *Proceedings of the 2025 ACM Conference on International Computing Education Research*, V.1. <https://doi.org/10.1145/3702652.374423>

1

Generative Artificial Intelligence in Academic Writing: A Critical Review of Scholarly Perspectives, Pedagogical Implications, and Emerging Governance Frameworks

- Ansari, S. (2026). Compound deception in elite peer review: A failure mode taxonomy of 100 fabricated citations at NeurIPS 2025. arXiv preprint arXiv:2602.05930.
- Aydin, O., Karaarslan, E., Erenay, F. S., & Bacanin, N. (2025). Generative AI in academic writing: A comparison of DeepSeek, Qwen, ChatGPT, Gemini, Llama, Mistral, and Gemma. arXiv preprint arXiv:2503.04765.
- Carobene, A., & Tang, E. (2024). Transparency and accountability in LLM-assisted scholarly manuscripts. *Journal of Research Ethics*, 12(1), 88–104.
- Chaka, C. (2024). AI detection reliability in multilingual writing contexts. *Language Assessment Quarterly*, 21(3), 211–230.
- Chanpradit, S. (2025). Gains and risks of generative AI in student writing: A systematic review of 30 empirical studies (2023–2024). *Educational Technology Research and Development*, 73(1), 1–28.
- Chen, Z., Wang, Y., & Liu, R. (2025). Impact of generative AI interaction and output quality on university students' learning outcomes. *Scientific Reports*, 15, 18024. <https://doi.org/10.1038/s41598-025-08697-6>
- Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 4, 100118. <https://doi.org/10.1016/j.caeai.2022.100118>
- Chiu, T. K. F. (2024). A classification tool to foster self-regulated learning with generative artificial intelligence by applying self-determination theory: A case of ChatGPT. *Educational Technology Research and Development*, 72, 2401–2416. <https://doi.org/10.1007/s11423-024-10366-w>
- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228–239. <https://doi.org/10.1080/14703297.2023.2190148>
- Diaz, C., & colleagues. (2025). ChatGPT in K–12 STEM classrooms: Perceived gains in interest, proficiency, and independence. *Journal of Science Education and Technology*, 34(1), 22–39.
- Digital Education Council. (2024). Global student AI usage survey 2024. Digital Education Council Publications.
- Finkel-Gates, A. (2025). ChatGPT in academic assessments: Upholding integrity. *Journal of Learning Development in Higher Education*, 36. <https://doi.org/10.47408/jldhe.vi36.1491>
- Floridi, L. (2023). Generative AI: Challenges for academic integrity. *AI & Society*, 38(4), 1–5.
- Freedom for All Americans. (2025, November 12). The ongoing battle in higher education over using ChatGPT for academic work. <https://freedomforallamericans.org/chatgpt-education-debate/>

Generative Artificial Intelligence in Academic Writing: A Critical Review of Scholarly Perspectives, Pedagogical Implications, and Emerging Governance Frameworks

- Frontiers in Education. (2025). The impact of generative AI on academic reading and writing: A synthesis of recent evidence (2023–2025). *Frontiers in Education*, 10, 1711718.
<https://doi.org/10.3389/feduc.2025.1711718>
- Frontiers in Education. (2026). A systematic critical review of generative AI's impact on authorship, pedagogy, and integrity (2023–2025). *Frontiers in Education*, 11, 1769680.
<https://doi.org/10.3389/feduc.2026.1769680>
- Heidt, J. (2024). Generative AI as an assistive technology for disabled and non-native speaking students. *Disability & Education International*, 26(1), 55–72.
- Hosseini, M., Horbach, S. P. J. M., Holmes, K., & Ross-Hellauer, T. (2024). Open science at the generative AI turn: An exploratory analysis of challenges and opportunities. *Quantitative Science Studies*, 5(3), 600–620.
- International Conference on Learning Representations. (2025). ICLR 2026 LLM use policy. ICLR Blog.
<https://blog.iclr.cc/2025/>
- Jain, R., & Jain, A. (2023). Generative AI in writing research papers: A new type of algorithmic bias and uncertainty in scholarly work. *arXiv preprint arXiv:2312.10057*.
- Kalhor, A., & colleagues. (2025). Algorithmic bias and knowledge representation in LLM-generated citation networks. *Scientometrics*, 130(4), 2115–2138.
- Kasneci, E., Sessel, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274.
<https://doi.org/10.1016/j.lindif.2023.102274>
- Kim, J., Park, S., & Lee, H. (2025). Students' roles and concerns in GenAI-supported academic writing. *Computers & Education*, 210, 105098.
- Kohnke, L. (2024). The digital divide, premium AI access, and academic equity. *Journal of Educational Computing Research*, 62(5), 1002–1020.
- Kovari, A. (2025). Ethical use of ChatGPT in education: Best practices to combat AI-induced plagiarism. *Frontiers in Education*, 9, 1465703.
<https://doi.org/10.3389/feduc.2024.1465703>
- Kosmyna, N., & colleagues. (2025). Neural and behavioural consequences of LLM-assisted essay writing: A laboratory experiment. *Nature Human Behaviour*, 9, 45–59.
- Lamba, N., Tiwari, S., & Gaur, M. (2025). Hallucinations in scholarly LLMs: A conceptual overview and practical implications. *Open Conference Proceedings*, 8.
<https://doi.org/10.52825/ocp.v8i.3175>

Generative Artificial Intelligence in Academic Writing: A Critical Review of Scholarly Perspectives, Pedagogical Implications, and Emerging Governance Frameworks

- Marzuki, R., & colleagues. (2023). Disparities in AI access and learning outcomes in low-resource settings. *International Journal of Educational Development*, 99, 102800. <https://doi.org/10.1007/s13347-026-01068-7>
- Meyer, J., Jia, P., Rathner, E., & colleagues. (2023). ChatGPT and large language models in education: A review of opportunities and challenges. *MedEdPublish*, 12, 76. <https://doi.org/10.1038/s41599-026-06903-y>
- Nguyen, T. (2025). The epistemic downside of using LLM-based generative AI in academic writing. *Publications*, 13(4), 63. <https://doi.org/10.3390/publications13040063>
- Organisation for Economic Co-operation and Development. (2023). *OECD digital education outlook 2023: Towards an effective digital education ecosystem*. OECD Publishing. <https://doi.org/10.1787/c74f03de-en>
- Organisation for Economic Co-operation and Development, & European Commission. (2025). *Empowering learners for the age of AI: An AI literacy framework for primary and secondary education (Review Draft)*. OECD/EC.
- Springer Nature. (2025). Generative AI's impact on graduate student professional writing productivity and quality. *International Journal of Artificial Intelligence in Education*. <https://doi.org/10.1007/s40593-025-00528-z>
- Springer Nature. (2026). The new editorial gatekeepers: Understanding LLM-based interfaces, their benefits, risks and design. *Philosophy & Technology*. <https://doi.org/10.1007/s13347-026-01068-7>
- Springer Nature. (2026). Generative AI technologies and educational outcomes: A comprehensive meta-analysis comparing traditional and AI-driven approaches. *Humanities and Social Sciences Communications*. <https://doi.org/10.1038/s41599-026-06903-y>
- Tandfonline. (2025). Mastering knowledge: The impact of generative AI on student learning outcomes. *Studies in Higher Education*, 50(5), 714–735. <https://doi.org/10.1080/03075079.2025.2487570>
- UNESCO. (2023). *Guidance for generative AI in education and research*. UNESCO Publishing.
- van Niekerk, J., Delpont, P. M. J., & Sutherland, I. (2025). Addressing the use of generative AI in academic writing. *Computers and Education: Artificial Intelligence*, 8, 100342. <https://doi.org/10.1016/j.caeai.2024.100342>
- World Association of Medical Editors. (2023). Chatbots, generative AI, and scholarly manuscripts. *WAME Recommendations*. <http://www.wame.org/>
- Zhang, Y., Liang, C., & colleagues. (2024). ChatGPT use among medical students in the United States: Patterns, purposes, and concerns. *Medical Education Online*, 29(1), 2342905.

Generative Artificial Intelligence in Academic Writing: A Critical Review of Scholarly Perspectives, Pedagogical Implications, and Emerging Governance Frameworks