Original Article

Authenticity of Artificial Intelligence (AI) as a Citation Tool in Institutions of Higher Education: Review of Literature

Annerose Wanjiku Wang'ang'a, PhD[1]*

[1] St. Paul's University, Private Bag 00217, Limuru, Kenya.
[*] Author for Correspondence Email: rosemake158@gmail.com

ABSTRACT

The swift advancement of AI tools such as ChatGPT, Gemini, and Claude has ignited significant debate within learning facilities regarding their appropriate function in reviews in higher education. Although these tools present innovative opportunities for enhancing learning, creativity, and the exploration of knowledge, they are frequently misused as replacements for traditional search engines. Unlike established platforms like Google or Bing, which continuously update by indexing real-time web content, AI systems generate responses based on pre-existing, static training data. Thus, the data produced may be outdated, unverifiable, and lacking in clear source attribution, raising important concerns around reliability and clarity. In scholarly settings, researchers typically rely on search engines and scholarly databases that provide direct access to abstracts, references, and peer-reviewed literature. Conversely, AI-generated content may mimic scholarly writing but often lacks citations or includes unverifiable claims. A particularly troubling issue is known as "AI hallucination," where AI tools may invent data or references, giving the illusion of authority without factual grounding. Because of these issues, this paper emphasises that AI should not be viewed as a credible or primary source for academic evidence. Rather, such tools should serve a supportive role, helping users generate ideas, summarise complex texts, or clarify difficult concepts. To ensure responsible integration of AI in scholarly applications, academic facilities must create clear protocols and offer training that enables critical analysis of AI outputs, validation of information, and ethical usage. This article aims to explore the authenticity of AI as a citation tool and its limitations in higher education. In conclusion, this discussion promotes a careful, informed utilisation of artificial intelligence in scholarly contexts. With proper surveillance and a clear understanding of its limitations, AI can complement traditional investigative methods while preserving rigour in addition to the trustworthiness of academic scholarship.

157 |

INTRODUCTION

Artificial Intelligence (AI), a concept formally introduced at the Dartmouth Conference in 1956 by John McCarthy, aimed to create machines capable of mimicking human cognitive functions (Cohen-Almagor, 2011; Avicena Tech Corp, 2024). AI encompasses diverse disciplines such as statistical analysis, computational neurobiology, mathematical logic, and computer science. While early technological discourse largely revolved around the internet and computers, AI was not initially regarded as a central tool in academic or technical contexts.

A major leap in AI's development came with the advent of Generative Pre-trained Transformers (GPTs), which perform various linguistic tasks ranging from answering questions and generating essays to composing poems and engaging in dialogue. These capabilities have positioned AI tools like GPT at the forefront of contemporary computational innovation.

Search engines were developed to meet the growing need for organised retrieval of digital content, evolving alongside the internet and the proliferation of websites. Their efficiency has made them indispensable to media organisations, businesses, and educational institutions striving for high visibility online (Escandell-Poveda, Iglesias-Garcia & Papi-Galvez, 2022).

Lierbarman (2024) explains that computational intelligence first emerged in higher education through paraphrasing tools designed to mask copied content. These tools initially produced substandard output but have since evolved, leveraging generative AI and natural language processing to deliver higher-quality results. OpenAI's development of ChatGPT, a sophisticated, conversational AI system, has marked a milestone in this trajectory. Its extensive knowledge base enables it to generate articulate, academic-style responses, potentially rivalling traditional scholarly sources.

Sahar & Munawaroh (2025) assert that AI is reshaping higher education by automating routine activities like scoring and registration, including resource allocation, thereby enhancing operational efficiency. AI's flexible education tools can also individualise learning material, catering to individual learners' needs and progression.

This paper aims to critically evaluate the authenticity of AI-generated citations in academic contexts and to assess the implications of their use in higher education. It argues that while AI systems offer support in generating ideas and summarising content, their outputs should not be regarded as academically reliable sources due to concerns of accuracy, traceability, and scholarly rigour. The structure of the paper covers a literature-based methodology, evaluates AI's role in research

authenticity, examines hallucination risks, discusses ethical considerations, reviews guidelines, and ends with recommendations and a conclusion.

METHODOLOGY

This article employs a narrative literature review approach. Sources were identified through academic databases such as Scopus, Google Scholar, and Web of Science, using keywords including "AI in education," "ChatGPT citation," "AI hallucination," and "AI ethics in research." The review focused on literature published between 2018 and 2025, prioritising peer-reviewed journal articles, institutional guidelines, and policy briefs. Inclusion criteria centred on studies addressing AI's credibility, citation integrity, and ethical implications in academic settings. The analysis was organised thematically to highlight the current debates and evolving standards regarding AI use in scholarly communication.

Authenticity of Artificial Intelligence in Research

In scholarly research, search engines typically provide researchers with links to scholarly articles and abstracts upon entering a query. In contrast, AI tools like ChatGPT produce coherent, essay-like responses to similar prompts. Despite its eloquence, concerns remain about the reliability of ChatGPT's output, as it may include inaccuracies or fabricated content (Lieberman, 2024).

According to OpenAI (2022), ChatGPT may generate convincing but incorrect or nonsensical answers. With an estimated accuracy of 85%, it should be regarded more like a blog or corporate website than a scholarly source. Lanum (2023) cites Harvard Business School's Edward McFowland III, who warns against uncritically accepting ChatGPT's outputs, stressing that despite its conversational nature, it lacks true reasoning abilities. Nevertheless, using ChatGPT to draft an initial essay that guides further research can be a practical and beneficial approach.

A key concern identified across multiple studies (e.g., Anani et al., 2023; Ocen et al., 2022) is the tension between efficiency and critical thinking. While AI tools can rapidly generate summaries and responses, their use may discourage students from engaging in deeper cognitive processes required for analysis and synthesis.

Additionally, unlike search engines that offer transparency through links to academic sources, AI tools present synthesised knowledge without traceable citations. This lack of transparency creates a challenge in validating information and undermines the principles of academic integrity. Researchers are increasingly advised to cross-reference AI-generated insights with peer-reviewed sources to preserve scholarly accuracy.

Artificial Intelligence Hallucination

AI hallucination points to instances where AI systems generate plausible yet factually incorrect or misleading content. This phenomenon affects various AI applications, including chatbots like ChatGPT, image generators like DALL-E, and even autonomous vehicles (Choi & Mei, 2025). For large language models, hallucinations often manifest as inaccurate, fabricated, or irrelevant text.

Sun et al. (2024) classify hallucinated data into two categories: disinformation (intentional falsehoods) and misinformation (inaccuracies without malicious intent). These occur when AI misinterprets input or attempts to "fill in the blanks" based on incomplete or biased training data.

Despite mitigation efforts, users must remain cautious and verify AI-generated content through trusted sources or expert consultation. Choi and Mei (2025) stress the importance of vigilance in high-stakes areas like education, law, or medicine.

Farquhar and Kossen (2024) highlight a breakthrough at Oxford University, where researchers developed a statistical method to evaluate uncertainty in AI-generated content based on meaning rather than word sequences. By

converting predicted word probabilities into probabilities of meaning, this approach improves trust in AI-generated answers, especially in critical contexts.

One real-world example of AI hallucination occurred in 2023, when a U.S. lawyer cited nonexistent court cases in a legal brief prepared with ChatGPT. The fabricated references, though presented convincingly, were exposed during court proceedings, highlighting the dangers of overreliance without verification.

Ethical Considerations for Artificial Intelligence Use in Research

Malik et al. (2023) and Sahar and Munawaroh (2025) emphasise that while computational intelligence improves writing efficiency and academic performance, its responsible incorporation into higher education demands moral scrutiny, including issues on plagiarism, bias, and learning dependency.

Vieriu and Petrea (2025) caution against replacing human interaction with AI, noting that although AI offers efficiency and data insights, it lacks empathy and creativity. Overreliance can diminish opportunities for critical thinking and reduce reflective engagement in learning environments. Moral challenges like algorithmic bias, information concealment, and unequal access should be addressed to foster inclusive, equitable learning.

Anani, Nyamekye, and Koduah (2025), through their study on Ghanaian postgraduate students, observed that while AI tools improve writing efficiency, there is a risk of diminishing analytical skills and fostering overdependence. Similarly, Ocen et al. (2025) found that although AI supports scholarly writing, excessive reliance can undermine creativity while raising integrity issues, especially around data security and institutional compliance.

Butson and Smith (2024) note that AI's demand for large datasets creates ethical dilemmas. Informed consent, which is a core principle in traditional research, thus becomes complex when data can be reused unpredictably. Researchers must reexamine their integrity frameworks to ensure responsible AI deployment. Presently, ethical concerns regarding AI use in higher education are widely debated (Crowe et al., 2017; Pisica et al., 2023).

Guidelines on the Use of Artificial Intelligence in Research

The Academy of Science of South Africa (ASSAf Council) recommends ongoing updates to guidelines concerning AI use in scholarly communication. These guidelines should offer protocols for authors, editors, and publishers regarding citation practices. AI-generated content, like personal communication, should be referenced as unrecoverable. Authors should disclose AI usage in the methodology or acknowledgements section, as appropriate.

Laher (2025) and Veiga (2025) note that routine uses of AI, such as grammar correction or literature searches, do not require citation. However, any AI-generated content must be clearly justified, labelled, and verified for accuracy, bias, and plagiarism. Tasks like code debugging or figure generation may not be cited directly in the text but should be acknowledged in supplementary material.

Veiga (2025) also identifies a knowledge gap in ethical standards, calling for unified guidelines that distinguish between AI-assisted tasks and generative AI content. These guidelines should reinforce the human researcher's responsibility for ensuring honesty in scholarly work.

European Commission (2025) echoes a sentiment that highlights the significance of transparency in computational intelligence use, awareness of its challenges, like hallucinations and bias, and compliance with legal frameworks. Researchers must safeguard privacy, intellectual property rights, and authorship standards. Moreover, generative AI tools should not be used extensively in sensitive roles such as peer review or proposal evaluation. Continuous training in ethical AI use is encouraged.

Although the European Commission emphasises transparency and traceability in AI outputs, practical implementation within universities remains challenging. Various obstacles can arise, such as insufficient teacher expertise and inadequate technological infrastructure within schools (Celik et al., 2022). Additional concerns include issues of cost-efficiency, cultural incompatibilities (Fahimirad and Kotamjani, 2018), data privacy and security (Pisica et al., 2023), and fears of AI surpassing human control (Haseski, 2019).

Nevertheless, the integration of AI into education has sparked considerable debate due to emerging challenges and associated risks (Holmes and Miao, 2023). While AI tools offer university students the advantage of interactive and tailored learning experiences, higher education institutions, in particular, in middle-income nations, often struggle with limited financial and logistical support despite efforts to embrace innovation. This situation has sometimes motivated the development of affordable technologies designed to deliver customised assistance and services to students (Kuleto et al., 2021

Recommendations Based on the Current Literature,

To ensure the ethical and effective integration of artificial intelligence (AI) within academic research and practice, several key recommendations emerge. First, it is imperative that higher education institutions develop robust institutional policies that clearly articulate the acceptable use of AI in teaching, learning, and research contexts. These policies should delineate proper citation protocols, address ethical implications, and specify disciplinary procedures for misuse.

Second, comprehensive training for both students and faculty is essential. Institutions should implement educational initiatives aimed at enhancing digital literacy, with a particular focus on the critical evaluation of AI-generated outputs. Workshops and seminars that address AI ethics,

academic integrity, and scholarly writing conventions can play a central role in this effort.

Third, AI should be positioned as a complementary tool that supports, rather than replaces, traditional scholarly resources. Its use may be particularly beneficial in activities such as brainstorming, outlining, and elucidating complex concepts; however, it should not supplant engagement with primary academic literature.

Fourth, the principle of transparency must guide the use of AI in academic work. Researchers are encouraged to disclose any use of AI tools in the generation or analysis of content, thereby fostering accountability and maintaining trust in scholarly communication.

Finally, there is a pressing need for ongoing empirical research into the long-term implications of AI in educational settings. Future studies should explore its impact on student learning outcomes, ethical considerations, and the evolution of citation norms and academic standards.

CONCLUSION

In summary, while AI tools like ChatGPT, Gemini, and Claude offer transformative potential for enhancing academic engagement and knowledge exploration, their limitations as reliable sources of scholarly evidence cannot be overlooked. Their static training data, lack of real-time updates, and susceptibility to generating unverifiable or fabricated content underscore the risks of using them as citation tools in academic writing. Therefore, AI should be embraced not as a replacement for traditional scholarly resources but as a supplementary aid that supports idea generation, content clarification, and learning facilitation. For responsible and ethical integration into higher education, institutions must establish clear usage guidelines and equip learners and educators with the skills to critically assess AI-generated content. Only through such structured engagement can AI contribute meaningfully to the academic landscape without compromising the

integrity and credibility of scholarly work. Looking to the future, sustaining academic excellence in the age of artificial intelligence will depend on achieving a careful equilibrium between leveraging technological advancements and upholding human discernment. It is essential for universities and researchers to prioritise critical thinking, uphold transparency, and maintain ethical standards to ensure that AI serves to enhance rather than compromise the integrity and value of scholarly work.

REFERENCES

Anani, G.E., Nyamekye, E. & Koduah, D.B. (2025). Using Artificial Intelligence for Academic Writing in Higher Education: The Perspective of University Students in Ghana.Discover Education Research https://doi.org10.1007/544217-025-00434-5.

ASSAf Council (2024). ASSAF and SCIELO Guidelines for the Use of Artificial Intelligence (AI) Tools and Resources in Research Communications. Science and Innovation: Republic of South Africa.

Avicena Tech Corporation (2024). The History of Artificial Intelligence. https:www.avicena.tech.

Butson, R. & Smith, R.S. (2024). AI and Its Implications for Research in Higher Education: A Critical Dialogue pp 563-577 https://doi.org/10.1080/07294360.2023.2280200.

Celik, I., Dindar, M., Muukkonen, H., and Järvelä, S. (2022). The promises and challenges of artificial intelligence for teachers: a systematic review of research. TechTrends 66, 616–630. doi: 10.1007/s11528-022-00715-y

Chen, X., Zou, D., Xie, H., Cheng, G., Liu, C. (2022). Two Decades of Artificial Intelligence in Education: Contributors, Collaborations, Research, Topics, Challenges and Future Directions. EduTechnology Soc.25:1

Crowe, D., LaPierre, M., and Kebritchi, M. (2017). Knowledge based artificial augmentation intelligence technology: Next step in academic instructional tools for distance learning. TechTrends 61, 494–506. doi: 10.1007/s11528-017-0210-4

Choi, A. & Mei, K.X. (2025). What are AI Hallucinations? Why AIs Sometimes Make Things Up. Retrieved from https://theconversation.com/what-are-ai-hallucinations-why-ais-sometimes-make-things-up-242896.

CohenAlmagor, R. (2011). InternetHistory. International Journal Tech.https://doi.org/10.4018/jte.2011040104.

Escandell-Poveda, R., Iglesias-Garcia, M. & Papi-Galvez, N. (2022). From Memex to Google: The Origin and Evolution of Search Engines.

European Commission (2025). Living Guidelines on the Responsible Use of Generative AI in Research. The International Association of Scientific, Technical and Medical Publishers.Https://www.natre.com/articles/d41586-023-03266-1.

Fahimirad, M., and Kotamjani, S. S. (2018). A review on application of artificial intelligence in teaching and learning in educational contexts. Int. J. Educ. Dev. 8, 106–118. doi: 10.5296/ijld.v8i4.14057

Farquhar, S. & Kossen, J. (2024). Major Research into Hallucinating Generative Models Advances Reliability of Artificial Intelligence. University of Oxford.https://www.ox.ac.uk/news/2024-06-20.major-research-hallucinatiing-generative-models-advances-reliability-artificial.

Haseski, H. I. (2019). What do Turkish pre-service teachers think about artificial intelligence? Comput. Sci. Educ. Sch. 3, 3–23. doi: 10.21585/ijcses.v3i2.55

Kuleto, V., Ilić, M., Dumangiu, M., Ranković, M., Martins, O., Păun, D., et al. (2021). Exploring opportunities and challenges of artificial intelligence and machine learning in higher education institutions. Sustainability 13:10424. doi: 10.3390/su13181042

Laher, S. (2025). Can Academics Use AI To Write Journal Papers? What the Guidelines Say; University of Witwatersland.

Lanum, N. (2023). AI Experts. Professors Reveal How Chat GPT Will Radically Alter The Classroom:' Age of the Creator'Fox News Digital. https://www.foxnews.com/media/ai-experts-professors-chatgpt-radically-alter-classroom-age-creator.

Lierberman, G. (2024). The Use and Detection of AI-Based Tools in Higher Education.Journal Of Instructional Research, Volume 13:Grand Canyon University.

Lopez-Chila, R., Llerena-Izquierdo, J., Sumba-Nacipucha, N., Cueza-Estrada, J. (2024). Artificial Intelligence in Higher Education: An Analysis of Existing Biliometrics. Educ.Sci.https://doi.org/10.3390/edusci.4010047.

Malik, R., Pratiwi, Y., Andajani, K., Numertayasa, I.W./Suharti, S., Darwis, A. & Marzuki (2023). Exploring Artificial Intelligence in Academic Essay: Higher Education Students Perspective.https://doi.org/101016/jedro.2023.100296.

Ocen, S., Elasu, J., Aarakit, S.M. and Olupot, C. (2025). Artificial Intelligence in Higher Education Institutions: Review of Innovations, Opportunities and Challenges. Front.Educ.10.1530247.doi:10.3389/feduc.2025.1530247.

Open AI's Chat GPT and the Prospect of Limitless Information: A Conversation with Chat GPT(2023). Journal of International Affairs, Volume 75(1)pp379-385

Pisica, A. I., Edu, T., Zaharia, R. M., and Zaharia, R. (2023). Implementing artificial intelligence in higher education: pros and cons from the perspectives of academics. Societies 13:118. doi: 10.3390/soc13050118

Sahar, R. & Munawaroh, M. (2025). Artificial Intelligence in Higher Education with Bibliometric and Content Analysis for Further Agenda. Discov Sustain Volume 6 pp401.https://doi.org/10.1007/543621-025-01086-z

Sun, Y., Sheng, D., Zhou, Z. & Wu, Y. (2024). AI Hallucination: Towards a Comprehensive Classification of Distorted Information in Artificial in Artificial Intelligence Generated Context.Humanities and Social Sciences Communications, Volume II.

Veiga, A.D. (2025). Ethical Guidelines for The Use of Generative Artificial Intelligence and Artificial Intelligence Assisted Tools in Scholarly Publishing: A Thematic Analysis. Science Editing,Volume 12(1)pp28-34.Doi:https:/doi.org/10.6087/kcse.352.

Vierio, A.M. & Petrea, G. (2025). The Impact of Artificial Intelligence (AI) on Students' Academic Development.Education Sceinces Volume 15(3)pp343.https://doi.org/10.3390/edusci15030345.