The Second Bridge on Artificial Intelligence for Scholarly Communication (AAAI-26)
Contributions
https://doi.org/10.52825/ocp.v8i.3175

Published: 18 Dec. 2025

Hallucinations in Scholarly LLMs

A Conceptual Overview and Practical Implications

Naveen Lamba[1,*], Sanju Tiwari[1], and
Manas Gaur[2]

[1]CAIMIF & Dept. of CSA, Sharda University, India

[2]University of Maryland, Baltimore County, USA

*Correspondence: Naveen Lamba, naveenlamba30894@gmail.com

Abstract. The issue of large language models (LLMs) is gradually infiltrating the academic workflow, but it also presents one significant problem: hallucination. The hallucinations involve invented research results, ideas of fabricated reference, and misinterpreted inferences that destroy the credibility and dependability of scholarly writing. In the present paper, the concept of hallucinations as the aspect of scholarly communication is discussed, the major types of hallucinations are revealed, and the causes along with effects of hallucinations are discussed. It also examines pragmatic mitigation measures, such as retrieval-augmented generation (RAG) of factual grounding, citation-verification, and neurosymbolic strategies of structured fact-checking. The paper additionally emphasizes the significance of human-AI partnership in the process of creating scholarly tools to make the use of AI in research responsible and verifiable. The paper seeks to create awareness and offer guidance to the creation of reliable AI systems to be used in scholarly contexts by synthesizing risks, opportunities, and available mitigation measures to such systems. Instead of presenting a comprehensive technical structure, the work provides an overview of the conceptual description which may be used to design more reliable, transparent, and fact-driven AI-assisted research tools.

Keywords: Hallucination, Large Language Models, Knowledge Graphs, Scholarly Communication, Neurosymbolic AI, Retrieval-Augmented Generation

1. Introduction

The emergence of large language models (LLM) is reshaping the academic world rapidly, helping researchers to formulate hypotheses, review the scientific material, summarize the literature, and write papers [1], [2]. Although these models can simplify the academic processes and make academic knowledge more accessible, as their popularity increases, so does an intrinsic problem, which can be described as hallucination. There are hallucinations, when models generate confident and factually inaccurate, fabricated or unverifiable information [3], [4].

The problem of hallucinations in academic circles is a real danger as it may result in misread discoveries, nonexistent citation, and distorted bibliographic sources all of

which affect the integrity of academic work [5], [6], [7]. Some reports indicate that Bard and ChatGPT among others are prone to generating artificial references or even misattributing research in literature synthesis and systematic review activities [8], [9]. These errors may confuse the readers, undermine the quality of scientific products, and corrupt citation networks in the long run [10], [11]. Morally, the prevalence of the use of the material produced by the LLM under the guise of verifying their knowledge poses the threat of misinformation, prejudice, and unintentional academic dishonesty. These risks are even increased by their use in the assessment of literature and peer review wherein the models that are used are not factually based [2]. These issues reveal the necessity to ensure the balance between the fluency of language and the truth in AI-assisted academic writing [12], [13].

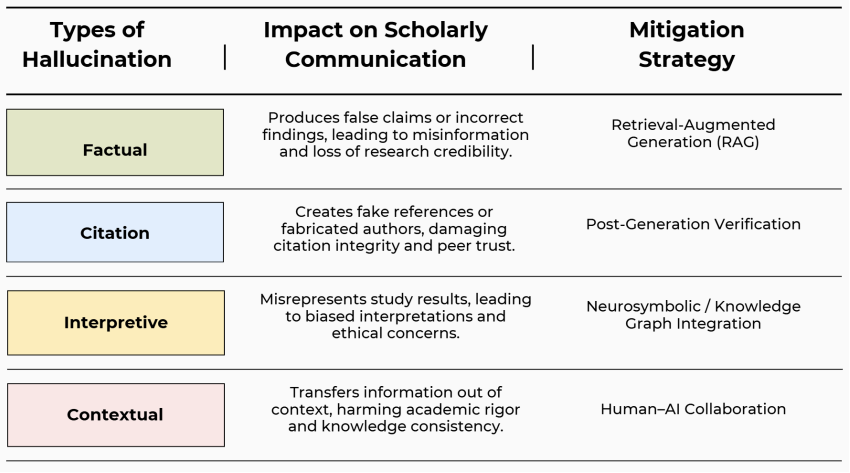

Figure 1. *The types of hallucinations mapped to their impacts and the mitigation strategies designed to address them*

The given paper summarizes the phenomenon of hallucinations in academic LLMs aiming to enhance the understanding of the issue and assist in the responsible implementation of AI in academia. We categorize the main forms of hallucinations, analyze their effects and generalize the most commonly talked of mitigation measures that may improve the factual reliability [14], [15]. The contribution of the work is threefold: (i) a brief conceptual synthesis of hallucination phenomena in scholarly publications and peer-review processes, (ii) an amalgamation of findings in recent surveys and empirical research in classifying the types of hallucination and their determinants, (iii) the description of possible practical ways to reduce these risks of those occurring through evidence-based approaches, hybrid symbolic-neural methods, and responsible human-AI cooperation. Figure 1 is the summary of the connection between the forms of hallucinations, their effects on scholarly communication, and the mitigation measures.

2. Methodology

In order to form an orderly insight into hallucinations in academic LLMs, this research paper uses a selective review method. Since the sphere is evolving quickly, and the contribution is conceptual, the aim of the methodology is to guarantee the transparency of the way the existing work was found, decoded and incorporated into the analytical grid of this paper.

Literature Identification: The relevant literature was determined with the help of the iterative search in the large scholar databases Google Scholar, Scopus, ACM Digital Library, SpringerLink and arXiv. Key search terms were combined to form the

search strategy of the following: LLM hallucination, citation hallucination, scientific misinformation, retrieval-augmented generation, knowledge graph fact-checking, and AI in scholarly communication. The output of this process includes publications between 2019-2025, both general literature on the behavior of LLM and the newer literature on hallucinations in academic settings.

Selection Criteria: The sources were filtered to provide relevance to the scope of this paper with focus to studies that: (i) study the behaviors of text-based LLMs on hallucinating, (ii) analyze the use of LLM in the academic or scientific processes, or (iii) suggest ways to mitigate risks in academic writing, citation generation or research synthesis. The inclusion of conceptual essays, empirical analyses, surveys, and technical reports was done in order to cover the thematic in a comprehensive manner, but only the works that are not related to scholarly communication or aim at multimodal hallucinations were excluded.

Analytical Approach: The selected works were examined in order to identify common themes in regard to the nature, causes, and effects of hallucinations in scientific communication. The analysis of insights was based on qualitative synthesis, which allows merging four dimensions of analysis, which structure this paper: (i) types of hallucinations, (ii) underlying causes, (iii) influences academic communication, and (iv) mitigation strategies. The conceptual mapping in Figure 1 was informed with this thematic integration as it connects the type of hallucination with the scholarly implications and the corresponding mitigation paths.

Scope and Limitation: Although such a methodology provides systematic knowledge, it is not a systematic review. The dynamic character of the research in the field of the LLM, as well as the scarcity of domain-specific benchmarks, entail the fact that the literature is still fragmented. Still, the approach facilitates a logical and scholarly-inspired synthesis that emphasizes the existing body of knowledge and outlines the new trends of enhancing reliability in academic LLMs.

3. Understanding Hallucination in Scholarly Context

Hallucination in the context of LLMs denotes text generated that seems coherent without being backed by the input data, factual evidence or knowledge base. More strictly speaking, it is fabrication, misrepresentation or mistaken recollection of information that distorts verifiable truth [3], [15]. Such misrepresentations may cause a loss of academic credibility in academic communication because the models can be used to create studies, falsely assign findings, or create or alter bibliographic information [1], [5]. The problems obscure the distinction between acceptable scientific synthesis and AI-produced harmful information, major ethical and epistemological issues in the academic community are brought up [16].

3.1 Types of Scholarly Hallucinations

In general, there are four main types of academic hallucinations, and each of them has a specific implication on scholarly communication [11], [17]:

Factual hallucination: It is a case where an LLM produces inaccurate, verifiable or fabricated data. As an example, a causal relationship between two variables is claimed when there is no such causal evidence in the literature..

Citation Hallucination: This is when a bibliographic element, like an author, article title, journal title, conference venue, publication date or DOI, is made up or modified.

LLMs like ChatGPT and Bard are known to produce fake-looking but nonexistent sources regularly, all of which are reported in the recent literature [6], [8], [9], and the integrity of academic citation practices is directly threatened.

Interpretive hallucination: A hallucination that is caused by a model malrepresenting, overgeneralizing, or misinterpreting the results of a source article. They may include drawing a causal conclusion based on a correlational finding or generalizing a research in a manner that goes against its findings.

Contextual Hallucination: This occurs when an LLM uses old, domain-irrelevant, or lost context to make predictions. To illustrate, the application of biological analogies to explain social science phenomena in an unjustified manner [7], [10].

The individual types are a symptom of failure by an LLM to bring generated language to bear on facts. The need to create sophisticated categorizations and effective measures of evaluation, like the new benchmark of hallucinations in scientific texts called SciHal25 [18], is also driven by these differences.

3.2 Why They Occur

LLMs are not fact-oriented and are more accurately described as being linguistic fluency oriented, thus giving them a better preference for coherent text over verifiable truth [3], [15]. They do not have real-time access to authoritative academic data, like PubMed, Google Scholar, or Scopus, and as such, the outputs obtained are out of date or unverifiable, due to their reliance on constant pretraining data [2], [14]. RAG systems can alleviate–although not entirely get rid of–hallucinations; when the evidence to be retrieved is irrelevant, noisy, or misaligned with the query presented by the user, they are still capable of generating false information [19], [20].

Overconfidence bias is another determinant of hallucinations, as in this case models give out responses with unwarranted confidence on instances where the internal probability distributions of the model show that they are uncertain. This is due to the fact that autoregressive token prediction can give out fluent statements that are incorrect- an even greater cause of concern in the academics environment [11].

4. Impact on Scholarly Communication

Academic workflows, including literature searches and manuscript writing, and initial support in peer review, are becoming more and more automated by LLMs. Although such systems are able to improve accessibility to information, and productive researcher work, their hallucinatory tendencies are causing large threats to the integrity and reliability of academic communication [5], [10].

4.1 Research Credibility

Academic study relies on the provable and traceable evidence. In cases where the LLMs forge references, they give the false impression that there is support in the form of references. According to research conducted by Chelli et al. [8] and Cheng et al. [9], 30-40 percent of the citations produced by models like ChatGPT and Bard are wrong or completely fake. To use transformer interpretability as an example, challenged to name major works on transformer interpretability, an LLM can list a few plausible-looking references, which cannot be found in any indexed database. The realistic formatting coupled with fake yet authoritative sounding names can deceive the researchers or

reviewers to accept such references as real ones. This erodes the credibility of academic writing that is supported by AI and makes peer review challenging.

4.2 Knowledge Propagation and Scholarly Memory

The information that is hallucinated does not exist in isolation. It will be able to spread once introduced to secondary analyses, review papers, educational content, or academic wiki. The inaccurate facts produced by LLMs can be included in literature reviews, meta-analyses, or other future AI-generated summaries, such that they contaminate the scientific record at large. According to Mustaffa et al. [11] and Nsirim et al. [10], the inaccuracy that is already present in such content may increase with the repetition of the same content, forming a cycle of misinformation. Gumaan et al. [15] calls this phenomenon the recursive contamination which is an echo-chamber of AI-generated text being fed back into training or summarization pipelines. By so doing, the effects of hallucinations not only jeopardize individual outputs but also the epistemic wholeness of the digital scholarship.

4.3 Integrity and Moral Accountability

Academic integrity is also an issue that arises because of hallucinations. Village The publication of the content generated by the LLM without proper verification can lead to unintended misrepresentation, academic dishonesty, or unintended plagiarism [16]. When researchers use fake references they jeopardize the effectiveness of the transparency and accountability of scholarly communication as well as undermine previous work. The paraphrases created by LLM can also be erroneous but subtly distort the meaning of the original findings yet linguistically persuasive [4], [12]. Teachers and publishers stress more the necessity of transparency and verification, claiming that ethical utilization of AI tools in researches can only be guaranteed with the assistance of human oversight [21].

4.4 Trust Crises and Reproducibility

The presence of hallucinations created by the LLMs endangers the principle of reproducibility, as it distorts the facts of empirical data. In the event that an LLM has inaccurately summarized results, or reported some form of methodological information, a future researcher will not have the capacity to replicate findings using incorrectly reported secondary descriptions. This is an example of a clinical trial that is summarized to exaggerate the size of the sample or an exaggerated effect, and meta-analyses based on the summary can produce false conclusions. According to King et al. [13], even though certain hallucinations do not cause any systematic bias, others induce systematic bias that kills trust between researchers and AI systems. These distortions are a reflection of the greater reproducibility crisis witnessed throughout the scientific world, and it is reflected in AI-generated text.

5. Mitigation Strategies

Although hallucinations could jeopardize the integrity of scholarly communication, there are a number of strategies that have been advanced to assist in reducing their impact. In the following section, the most important mitigation strategies to enhance the level of factual grounding, accuracy of citation, and transparency of model reasoning in academic settings are reviewed [2], [3], [19].

5.1 Retrieval-Augmented Generation

One of the most common methods of minimizing hallucinations, as well as in the academic world, is retrieval-augmented generation (RAG). Instead of using pretrained parameters only, RAG-enabled models access the documents that are considered to be relevant within trusted sources, i.e. Google Scholar, Scopus, or the Open Research Knowledge Graph, to base their answers on actual evidence. This makes probabilistic inference less relied upon and assists in making sure that the outputs are based on verifiable publications.

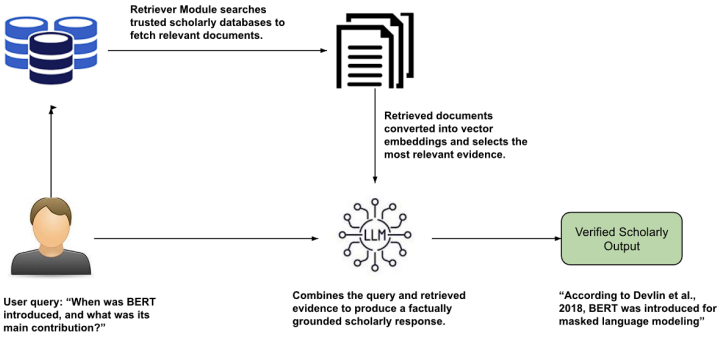

Figure 2. *Workflow of RAG*

In current extensions like RAG-HAT [20] hallucination-aware tuning pipelines are trained to incorporate retrieved evidence and are discouraged to make unsupported claims. This makes them very useful in academic writing assistants, automated literature review systems and citation recommendation systems because such frameworks allow users to trace generated claims or citations to verified sources. The figure 2 depicts a standard RAG process within academic communities, according to which a model identifies and encodes evidence based on reliable academic databases to minimize factual and citation based mistakes.

Next to conventional RAG pipelines, recent research focuses on the interpretability and the reliability of retrieval as such. METEORA [22] uses explicit reasoning paths to select dynamically the pertinent evidence instead of using opaque top-k selection heuristics, which enhances its resilience to misleading or adversarial agents. By incorporating such interpretability-oriented elements with RAG, it is possible to have more traceable, auditable, and reliable AI systems that can be used in scholarly research.

5.2 Post-Generation Verification

Hallucinations may still take place even when grounding is done by retrieval, and it is necessary to verify the post-generation. Checking systems may be in the form of rules or API-based, and allow the checking of the authenticity of citation, the correctness of bibliographic structure and compatibility with the authoritative databases. As an example, systems are able to cross-match journal titles, DOIs, year of publication and author names and then insert citations into generated outputs [7], [9].

According to Bareh et al. [5], semantic inconsistencies, e.g. misreported research results or misattributed methodology, should also be detected by verification layers by matching generated summaries to the entire text of the cited papers. Adding uncertainty or confidence scores to model outputs may also be useful in allowing the user to guide statements that are to be further verified [8]. These steps do not only enhance the

factual accuracy of the outputs of the LLM but also enhance responsibility in AI-based academic writing.

5.3 Neurosymbolic or Knowledge Graph Approaches

The next strategy is to incorporate neurosymbolic AI in the academic LLMs by merging the generative behavior of neural models with logical reasoning properties of symbolic systems [15]. Knowledge graphs are formalisations of objects, articles, datasets, writers as well as techniques and offer a factual foundation to confirm or limit created assertions [2].

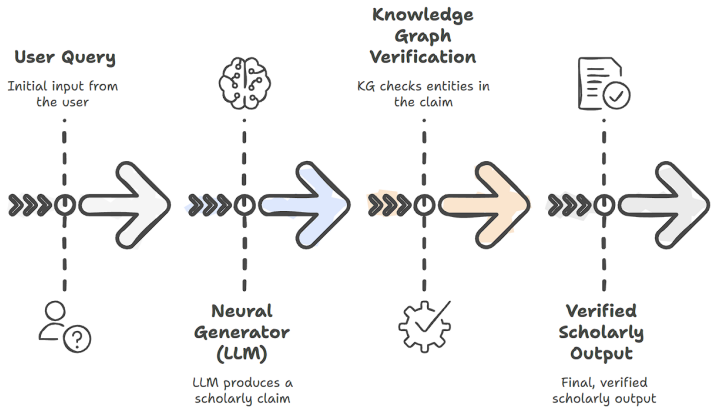

Figure 3. *Neurosymbolic framework for fact verification in scholarly LLMs*

Considering the next example, when an LLM claims that BERT was published in 2019, one may query a knowledge graph and verify the right publication year and source. Similarly, when a model says that study proved X, the system can check whether it is reflected in the metadata of the related paper or the citation graph [11]. Neurosymbolic integration can therefore act as a real-time fact-checking layer, which can be used to make LLMs more verifiable reasoning assistants instead of probabilistic generators. The neurosymbolic framework as Figure 3 shows, confirms model-generated claims by aligning them with entities and relations in a scholarly knowledge graph, allowing factual errors to be corrected, say by changing the year of publication or changing authors.

Nevertheless, there are shortcomings of the current scholarly knowledge graphs: they do not always have complete coverage, they are slow to index new works, and they would usually only record the metadata, but not the detailed methodological or interpretive information. Such gaps limit the capacity of verifying the complex scholarly claims, and neurosymbolic approaches to the combination of structured facts and contextual reasoning are required [23].

5.4 Human-AI Collaboration

With the advancement in technology, hallucinations cannot be completely done away with without the intervention of humans [1], [16]. The content produced by LLM must therefore be regarded as an assistant draft but not as a finished academic writing. Human-AI cooperation allows the iterative improvement, where the user validates citations, verifies interpretations and gives corrective feedback that can guide further improvements in the model [21].

Li et al. [18] propose the addition of prompts and checkpoints to hallucination-sensitive and user verifiable scholarly tools. These systems may indicate possibly

unsubstantiated claims, emphasize statements that need to be verified, or they may even contain links to legitimate sources. This practice fosters openness and the significance of human factor in the research processes with the support of AI.

6. Discussion and Future Directions

The application of LLMs in academic communication has some great opportunities and also challenges. Although these models promise in such activities like literature synthesis and initial peer-review assistance, their hallucinatory nature jeopardizes academic validity and repeatability. In order to improve the research in this field, it is necessary to create domain-specific datasets and evaluation benchmarks. Even though the earlier work like SciHal25 [18] establishes tasks to detect hallucinations in scientific text, more general and specific datasets are required to see hallucination behaviors in non-scientific areas. This would allow gradual evaluation of the reliability of LLM in citation generation, literature summary, hypothesis formulation and other educational activities.

The next generation systems must be able to include explainability processes that do not only indicate the possibility of hallucinations, but also explain why a particular output can be unreliable. According to a number of studies, this type of transparency can be used to reduce the gap between the black-box neural generation and human reasoning and make the decisions of the models easier to understand. Traceability and verification can also be supported by connecting generated claims with verifiable entities in structured knowledge graphs. Furthermore, incorporation of hallucination-detection algorithms into academic peer-review and publication pipeline might be helpful to the editors, as it allows recognizing incorrect paraphrasing, wrong references, or unsubstantiated arguments prior to publication. This would turn AI into an active, not a passive, text generator, or the so-called trust co-pilot, which would enhance integrity and not add noise. In the end, it will be aimed at creating reliable, clear, and hallucination-conscious AI systems that facilitate instead of corrupt scholarly communication.

7. Conclusion

In this paper, the conceptual perspective of the LLM-based hallucinations was presented in the framework of scholarly communication. It categorized core forms of hallucinations, explored the causes of such hallucinations, and discussed the consequences of such hallucinations on the academic integrity, academic reproducibility and academic trust. Among the mitigation measures, retrieval-augmented generation (RAG), post-generation verification, and neurosymbolic integration are discussed as they are intended to enhance the factual accuracy of the output of the LLMs. With the increased adoption of AI tools in academic processes, there will be a need to balance automation and accountability. To meet the criterion of verifiability and transparency, they will need domain targeted datasets, systems mindful of hallucinations, and explainable detection models. Finally, the missing component to turn LLMs into reliable research assistants thinkable, transparent, verifiable, and trustworthy AI helpers is the ability to see these technologies as more than mere text generators and instead as sources of strong support to scientific knowledge.

Competing interests

The authors declare that they have no competing interests.

Author contributions

Naveen Lamba: Writing – original draft; Writing – review & editing. Sanju Tiwari: Supervision. Manas Gaur: Supervision.

References

[1] J. G. Meyer et al., "Chatgpt and large language models in academia: Opportunities and challenges", *BioData Mining*, vol. 16, no. 1, pp. 1–11, 2023. DOI: 10.1186/S13040-023-00339-9. [Online]. Available: https://link.springer.com/article/10.1186/s13040-023-00339-9.

[2] Y. Sun, D. Sheng, Z. Zhou, and Y. Wu, "Ai hallucination: Towards a comprehensive classification of distorted information in artificial intelligence-generated content", *Humanities and Social Sciences Communications*, vol. 11, no. 1, pp. 1–14, 2024. DOI: 10.1057/s41599-024-03811-x. [Online]. Available: https://www.nature.com/articles/s41599-024-03811-x.

[3] S. M. T. I. Tonmoy et al., "A comprehensive survey of hallucination mitigation techniques in large language models", *arXiv preprint arXiv:2401.01313*, 2024. [Online]. Available: https://arxiv.org/pdf/2401.01313.

[4] Z. Bai et al., "Hallucination of multimodal large language models: A survey", *arXiv preprint arXiv:2404.18930*, 2024. [Online]. Available: https://arxiv.org/pdf/2404.18930.

[5] C. K. Bareh, "A qualitative assessment of the accuracy of ai-llm in academic research", *AI and Ethics*, vol. 5, no. 4, pp. 4305–4324, 2025. DOI: 10.1007/S43681-025-00730-8. [Online]. Available: https://link.springer.com/article/10.1007/s43681-025-00730-8.

[6] A. Jain, P. Nimonkar, and P. Jadhav, "Citation integrity in the age of ai: Evaluating the risks of reference hallucination in maxillofacial literature", *Journal of Cranio-Maxillofacial Surgery*, vol. 53, no. 10, pp. 1871–1872, 2025. DOI: 10.1016/J.JCMS.2025.08.004.

[7] J. Niimi, "Hallucinations in bibliographic recommendation: Citation frequency as a proxy for training data redundancy", *arXiv preprint arXiv:2510.25378*, 2025. [Online]. Available: https://arxiv.org/pdf/2510.25378.

[8] M. Chelli et al., "Hallucination rates and reference accuracy of chatgpt and bard for systematic reviews: Comparative analysis", *Journal of Medical Internet Research*, vol. 26, no. 1, e53164, 2024. DOI: 10.2196/53164. [Online]. Available: https://www.jmir.org/2024/1/e53164.

[9] A. Cheng, V. Nagesh, S. Eller, V. Grant, and Y. Lin, "Exploring ai hallucinations of chatgpt: Reference accuracy and citation relevance of chatgpt models and training conditions", *Simulation in Healthcare*, 2025. DOI: 10.1097/SIH.0000000000000877. [Online]. Available: https://journals.lww.com/simulationinhealthcare/fulltext/9900/exploring_ai_hallucinations_of_chatgpt__reference.198.aspx.

[10] O. Nsirim, "Hallucinations in artificial intelligence and human misinformation: Librarians' perspectives on implications for scholarly publication", *Folia Toruniensia*, vol. 25, pp. 79–98, 2025. DOI: 10.12775/FT.2025.004. [Online]. Available: https://apcz.umk.pl/FT/article/view/60574.

[11] N. E. Mustaffa, K. E. Lai, C. N. Preece, and F. Y. Wong, "A bibliometric review of large language model hallucination", *International Journal of Research and Innovation in Social Science*, vol. IX, no. 9, pp. 5025–5037, 2025. DOI: 10.47772/IJRISS.2025.909000409. [Online]. Available: https://rsisinternational.org/journals/ijriss/articles/a-bibliometric-review-of-large-language-model-hallucination/.

[12] J. Jamaluddin, N. Abd Gaffar, and N. S. S. Din, "Hallucination: A key challenge to artificial intelligence-generated writing", *Malaysian Family Physician*, vol. 18, p. 68, 2023. DOI: 10.51866/LTE.527. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC10726751/.

[13] M. R. King, "An update on ai hallucinations: Not as bad as you remember or as you've been told", *Cellular and Molecular Bioengineering*, pp. 1–6, 2025. DOI: 10.1007/S12195-025-00874-X. [Online]. Available: https://link.springer.com/article/10.1007/s12195-025-00874-x.

[14] S. Gupta, "Retrieval-augmented generation and hallucination in large language models: A scholarly overview", *Scholars Journal of Engineering and Technology*, vol. 13, no. 5, pp. 328–330, 2025. DOI: 10.36347/sjet.2025.v13i05.003.

[15] E. Gumaan, "Theoretical foundations and mitigation of hallucination in large language models", *arXiv preprint arXiv:2507.22915*, 2025. [Online]. Available: https://arxiv.org/pdf/2507.22915.

[16] A. Guleria, K. Krishan, V. Sharma, and T. Kanchan, "Chatgpt: Ethical concerns and challenges in academics and research", *Journal of Infection in Developing Countries*, vol. 17, no. 9, pp. 1292–1299, 2023. DOI: 10.3855/JIDC.18738.

[17] Z. Sun, "Large language models in peer review: Challenges and opportunities", *Scientometrics*, pp. 1–44, 2025. DOI: 10.1007/S11192-025-05440-W. [Online]. Available: https://link.springer.com/article/10.1007/s11192-025-05440-w.

[18] D. Li et al., "Overview of the scihal25 shared task on hallucination detection for scientific content", in *Proceedings of the 2025 Shared Tasks on Scientific Document Processing*, Association for Computational Linguistics (ACL), 2025, pp. 307–315. DOI: 10.18653/V1/2025.SDP-1.29. [Online]. Available: https://aclanthology.org/2025.sdp-1.29/.

[19] W. Zhang and J. Zhang, "Hallucination mitigation for retrieval-augmented large language models: A review", *Mathematics*, vol. 13, no. 5, p. 856, 2025. DOI: 10.3390/MATH13050856. [Online]. Available: https://www.mdpi.com/2227-7390/13/5/856.

[20] J. Song et al., "Rag-hat: A hallucination-aware tuning pipeline for llm in retrieval-augmented generation", in *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (Industry Track)*, Association for Computational Linguistics (ACL), 2024, pp. 1548–1558. DOI: 10.18653/V1/2024.EMNLP-INDUSTRY.113. [Online]. Available: https://aclanthology.org/2024.emnlp-industry.113/.

[21] V. Magesh, F. Surani, M. Dahl, M. Suzgun, C. D. Manning, and D. E. Ho, "Hallucination-free? assessing the reliability of leading ai legal research tools", *Journal of Empirical Legal Studies*, vol. 22, no. 2, pp. 216–242, 2025. DOI: 10.1111/JELS.12413.

[22] Y. Saxena, A. Padia, M. S. Chaudhary, K. Gunaratna, S. Parthasarathy, and M. Gaur, "Ranking-free rag: Replacing re-ranking with selection in rag for sensitive domains", *arXiv preprint arXiv:2505.16014*, 2025. [Online]. Available: https://arxiv.org/abs/2505.16014.

[23] T. R. Besold, A. S. d'Avila Garcez, L. C. Lamb, et al., *Neurosymbolic ai for egi: Aaai tutorial series*, https://nesy-egi.github.io/, Accessed: 2025-11-14, 2025.