International Journal of Teaching, Learning and Education (IJTLE)
ISSN: 2583-4371
Vol-4, Issue-6, Nov-Dec 2025
Journal Home Page: https://ijtle.com/
Journal DOI: 10.22161/ijtle

Hallucinations in Large Language Models for Education: Challenges and Mitigation

Kamaluddeen Usman Danyaro*[1], Shamsu Abdullahi[1], Abdallah Saleh Abdallah[2], Haruna Chiroma[3]

[1]Department of Computing, Faculty of Science, Management and Computing, Universiti Teknologi PETRONAS, 32610 Seri Iskandar, Perak, Malaysia
[2]Department of Arabic Language, Faculty of Islamic Studies, Universiti Islam Antarabangsa Tuanku Syed Sirajuddin (UniSIRAJ), Malaysia
[3]College of Computer Science and Engineering, University of Hafr Batin, Saudi Arabia
*kamaluddeen.usman@utp.edu.my

Received: 29 Sep 2025, Received in revised form: 25 Oct 2025, Accepted: 03 Nov 2025, Available online: 06 Nov 2025

Abstract

Large Language Models (LLMs) are increasingly being adopted in education to support teaching, learning, and assessment. While they offer benefits such as personalised learning and automated feedback, their tendency to generate hallucinations (plausible but factually incorrect or fabricated information) poses a critical challenge. In an educational context, hallucinations risk misleading students, compromising academic integrity, and eroding trust in AI-assisted learning. This paper examines hallucinations in education, highlighting their causes, risks, and implications. Unlike prior surveys that address hallucinations broadly, our work focuses specifically on education, where the consequences extend to academic honesty, critical thinking and equitable access. We provide a domain-specific analysis of how hallucinations emerge in tutoring systems, assessment and instructional content. Furthermore, we review technical and pedagogical mitigation strategies, such as prompt engineering, fine-tuning, dynamic course content integration and redesigned assessment practices. The paper contributes a framework that links technical solutions with education safeguards, emphasising that mitigating hallucinations is not limited to algorithmic advances but also requires institutional policies and critical AI literacy. By addressing these challenges, we aim to inform more reliable, equitable and trustworthy deployment of LLMs in education.

Keywords— Large Language Models (LLMs), Hallucination, Prompt Engineering, LLM for Education

I. INTRODUCTION

Large Language Models (LLMs) are transforming education by introducing new possibilities for personalised learning, task automation, and intelligent assistance (Sharma et al., 2025). These advancements allow students to benefit from adaptive tutoring and real-time feedback, while enabling educators to save time through the automated creation of instructional materials and assessments (Pirjan & PETROŞANU, 2024). By responding to individual learning needs, LLMs provide detailed explanations and promote inclusivity within the classroom (Lopez-Gazpio, 2025). This adaptability positions LLMs as partners that can work alongside teachers to enhance how education is delivered (Shahzad et al., 2025). For instance, they can lower educational barriers, create personalised learning paths and give students immediate constructive feedback (Razafinirina et al., 2024). Such adaptive learning models continuously assess learners' pace, strength and style, adjusting content and guidance to maximize understanding (Lopez-Gazpio, 2025). At the same time, educators gain valuable support, as LLMs handle pedagogical tasks such as generating instructional materials, creating assessments, and providing feedback, which increases teaching efficiency (Attard & Dingli, 2024). Beyond these practical benefits, LLMs also open new opportunities in foreign language education by offering immersive practice in speaking,

listening, reading and writing (Cherednichenko et al., 2024).

Despite these benefits, LLMs' adoption introduces critical risks. Foremost among these is hallucination (the generation of content that is inaccurate, fabricated or misleading) (Tonmoy et al., 2024). In educational contexts, such outputs can misinform learners, compromise academic integrity, erode trust in AI-assisted systems, and hinder the cultivation of critical thinking skills (Elsayed, 2024). These risks have been further exacerbated by training data biases, privacy concerns and the danger of excessive dependency on AI. To address these challenges, researchers emphasise the need for fairness, transparency and strong mitigation strategies. The practical solutions include bias reduction techniques, development of evaluation metrics to discover model limitations, and structured training programs to help teachers integrate LLM effectively into education. It is equally important to ensure that LLM is used in a manner that respects academic integrity, encourages independent critical thinking, and promotes responsible use (Nazi & Peng, 2024). While previous surveys have investigated hallucinations in a broader LLMs context, no study has systematically analysed them in educational-based LLMs. This paper addresses this gap by examining the causes and consequences of hallucinations in education. It also reviews technical and pedagogical strategies for mitigation, contributing to efforts to ensure that AI-powered learning remains reliable, equitable, and trustworthy. We believe that effective solutions must combine technological innovation and pedagogical safeguards to ensure that AI does not weaken learning but strengthens it.

The organisation of this paper is as follows: Section 2 provides an overview of LLMs' hallucination and its causes. Section 3 describes the implications for LLMs' Hallucinations in Education. Section 4 presents the hallucination mitigation Strategies. In Section 5, we conclude the study.

II. HALLUCINATIONS IN LLMS

Hallucination refers to instances where an LLM generates outputs that are factually incorrect or not grounded in real-world knowledge (Tonmoy et al., 2024). Such responses are often delivered with fluency and confidence, making them difficult for non-expert users to distinguish from accurate information. Although the metaphor is compelling, it is somewhat misleading: human hallucinations reflect distortions of perception, whereas LLM hallucinations stem from the probabilistic nature of language modelling. LLMs do not know facts but predict the most likely sequence of tokens based on training data (Fang et al., 2024). Hallucinations, therefore, are not malfunctions but inherent consequences of this design. When faced with gaps in knowledge, the model generates plausible but fabricated responses rather than withholding output. Addressing this issue requires more than simple error correction; it demands strategic approaches to design, deployment, and use. Two main categories of hallucinations are commonly identified. Intrinsic hallucinations occur when an output directly contradicts information provided in the input or a given source (Ji et al., 2023). For example, a model misreports an equation solution despite the student's correct input; this constitutes an intrinsic hallucination. Extrinsic hallucinations, by contrast, occur when the generated output introduces information inconsistent with real-world facts and unverifiable from available sources (Cossio, 2025). An example is fabricating references to non-existent educational psychology studies.

Causes of LLMs Hallucinations

Hallucinations in LLMs arise from multiple technical and design limitations. While these issues are common across domains, their effects are particularly problematic in education, where accuracy, fairness, and integrity are essential.

- **Training Data Limitations and Biases:** LLM reliability depends heavily on the quality of training data (Tonmoy et al., 2024). Educational risks arise because these datasets contain errors, outdated information, and cultural prejudices (Naser, 2025). For example, a model primarily trained on English-language sources can provide more accurate feedback to English-speaking native speakers when interpreting students' essays in low-resource languages. Likewise, conflicts with historical report on training data can lead to creating misleading explanations for classroom use. This could strengthen inequality and introduce misinformation into academic work.
- **Model Architecture and Probabilistic Nature**: LLMs are inherently probabilistic, generating text by selecting the most likely next token based on learned patterns (Mirchandani et al., 2023). This architecture prioritises fluency over factual accuracy, as models lack a genuine understanding of context. In education, this leads to problems when students step-by-step problem-solving model or an essay-writing model. For example, in mathematical reasoning, models can produce fluid but logically incorrect solutions. This is not due to a bug but because of the probabilistic design itself. For long-term assignments or research

Cross Ref DOI: https://dx.doi.org/10.22161/ijtle.4.6.2

tasks, the tendency of the model to fill gaps with plausible but false statements creates a risk of hallucination that students may struggle to detect.

- **Prompt Engineering Issues:** The clarity of user prompts has a significant impact on the LLM output (He et al., 2024). In educational settings, beginners often use vague or unspecified prompts, which causes the model to infer details incorrectly. For example, students who ask, "Explain Photosynthesis" without context may be very simplified or wrongly answered responses, while well-structured prompts (such as "Explain Photosynthesis for High School Students with Example Experiments") yield more accurate results. Because learners are not trained in rapid engineering, the danger of hallucinations caused by inaccurate prompts is particularly high in classrooms.
- **Transformer Limitations:** The transformer architecture that supports most LLMs limits memory and context. In education, this becomes a problem for long essays, dissertations, or extended teaching dialogues, because previous contexts may be lost, which leads to contradictions and incomplete answers. The tokenisation makes the problem even more complex: unusual academic terms or specific jargon in a discipline can be divided into sub-word units, distorting meaning. For example, special terms in chemistry or linguistics are misrepresented, which increases the likelihood of a hallucinatory or inaccurate explanation in the specific teaching of the subject.

Implications for LLMs Hallucinations in Education

LLM hallucinations pose significant challenges in education. While these models are increasingly adopted by students for benefits such as accessibility and task automation, their tendency to generate misleading information can undermine trust, impede learning outcomes, and emphasise the urgent need for AI literacy and improved model design. The subsequent discussion outlines the key implications of LLM hallucinations in educational contexts as described in Figure 1.

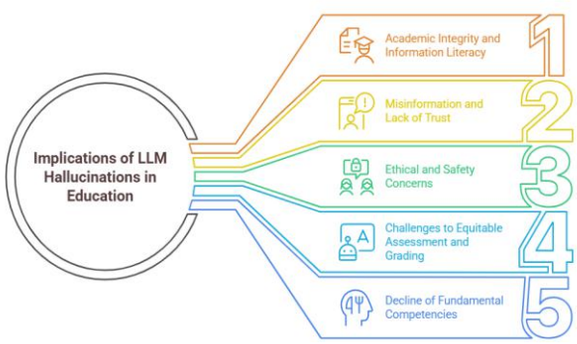

Fig.1. Implications of hallucinations in educational LLMs

Academic Integrity Risks

The integration of LLMs into the educational environment has triggered an increasing crisis of academic integrity and information literacy (Perkins, 2023). These models, which can write essays, solve exam questions and generate code, provide new ways of addressing academic ill-treatment by enabling students to outsource assignments, exams, or documents. Such abuse not only undermines the learning process but also creates an unfair advantage for those who rely on unethical artificial intelligence tools. The LLM often generates outputs containing false references, authors, or historical events, which makes it difficult for educators to verify the authenticity of their work and increases the risk of misinformation being embedded in academic work (Orenstrakh et al., 2024). The problem goes beyond fraud to fundamental threats to knowledge literacy. Students may struggle to determine whether the confident responses generated by LLMs are actually accurate, which leads to long-term misunderstanding, erodes critical thinking, and reduces motivation for learning. Efforts to combat this using AI detection tools such as GPTZero, Originality.AI, OpenAI text classification, and Turnitin's AI writing detection have shown mixed results. Research shows that even minor modifications, such as inserting a single word, significantly reduce the probability of detection and therefore undermine their reliability. In addition, these tools are at risk of creating false positives that affect non-English speakers in disproportionate quantities and raise concerns about digital inequalities and unjust accusations. Together, these challenges highlight the fact that LLM hallucinations and unidentified AI-generated content not only constitute a technical problem, but also an educational and ethical one, which requires educators to rethink evaluation design and encourage students to have greater digital literacy skills.

Misinformation and Lack of Trust

The main challenge of an LLM in education is that it tends to generate facts that are incorrect but persuasive (Elsayed, 2024). While the recent GPT-4 model has shown significant improvements over the previous version, hallucinations persist due to training data limitations and the probability generation method (Mohammed et al., 2024). LLM's own style often hides these inaccuracies and makes it difficult for students to evaluate the content critically. Consequently, learners may accept misleading or fraudulent information at face value, strengthening misperceptions instead of building real understanding (Elsayed, 2024). The consequences extend beyond individual learning. The persistence of hallucinations

Cross Ref DOI: https://dx.doi.org/10.22161/ijtle.4.6.2

reduces the trust of LLMs as educational tools and reduces their credibility among students and educators. For example, studies have shown that models such as GPT-4 sometimes generate citations, attribute content to non-existent authors, or provide irrelevant references to the subject (Toney, 2024). Such misinformation not only threatens knowledge literacy but also raises ethical concerns about the responsible integration of generative AI in education. These problems can prevent the adoption of LLM and limit its transformational potential in education and learning if not addressed.

Ethical and Safety Concerns

The black box characteristics of LLMs raise ethical and safety issues in education, especially their control and reliability (Cossio, 2025). Hallucinations can cause chatbots to deviate from their intended purpose, resulting in unintended or harmful results. These risks highlight the need for protective measures that ensure that LLMs meet educational objectives. With the rise of adoption, privacy, prejudice, and transparency issues are becoming more and more important. The protection of student data, equitable access and the prevention of algorithmic prejudices are the central elements of responsible use. To address these concerns, a clear ethical framework must be developed, which emphasises transparency in model training, the protection of sensitive information, and inclusiveness to reflect different student populations. Resolving these challenges is essential not only to mitigate risks but also to build the confidence needed for a sustainable integration of LLMs into education.

Challenges to Equitable Assessment and Grading

The use of AI in grading and assessment presents significant ethical and pedagogical challenges, many of which are heightened by LLM hallucinations. While such systems can ease teachers' workloads by flagging surface-level errors, they lack the human capacity to interpret context, recognise individual learning needs, and provide holistic feedback (Madsen et al., 2025). Hallucinations further undermine reliability, as models may generate fabricated justifications, misassign grades, or offer misleading guidance, eroding student trust in the assessment process. Algorithmic bias compounds these risks, particularly for students from diverse linguistic or cultural backgrounds who may be unfairly misjudged (Buolamwini & Gebru, 2018). Additionally, the high cost of implementing AI tools threatens to widen the digital divide, limiting equitable access (Selwyn, 2019). Ultimately, relying on AI as an arbiter of student performance risks shifting education from a process of growth and self-discovery toward a mechanistic exercise of aligning outputs with model preferences.

The Decline of Fundamental Competencies

The loss of fundamental abilities is one of the main risks associated with LLM hallucinations in the classroom (Elsayed, 2024). Students who rely too much on AI may have a "crutch effect," in which they are unable to internalise the skills required for deep learning (Gouscos). For instance, critical thinking is not innate; rather, it is acquired by challenging the facts, considering different viewpoints, and considering presumptions. This mechanism is directly undermined by hallucinations. LLMs lessen students' motivation to double-check assertions or conduct in-depth research by generating accurate but erroneous information. Furthermore, they avoid the intellectual conflict that promotes comprehension by responding in an authoritative, one-source manner. Therefore, hallucinations are more than just factual mistakes; they undermine the development of critical mental habits that education aims to foster, such as curiosity, skepticism, and analytical reasoning.

Hallucination Mitigation Strategies in Educational LLMs

Mitigating hallucinations in LLMs requires both technical solutions and pedagogical interventions. While technical methods focus on reducing factual errors during model generation, pedagogical strategies ensure that students and educators engage critically with AI outputs rather than accepting them uncritically.

Retrieval-Augmented Generation (RAG)

RAG is an important strategy for reducing hallucinations in LLMs, especially in education. Unlike traditional closed-book models, RAG produces reliable external sources such as academic databases, digital libraries and institutional data centres and converts LLM into an open book system. This process involves encoding documents into vector representations, extracting the most relevant materials in response to queries, and integrating them into prompts. This improves the accuracy of facts and ensures compatibility with verified educational knowledge (Bhattacharya, 2024). Advances such as Hyper-RAG have further enhanced this framework by recording higher-order relationships between documents extracted and reducing hallucinations (Feng et al., 2025). However, effectiveness depends on the quality of the educational repository: a lack of recovery or dependence on outdated sources still leads to misleading content, making careful observation a necessity for reliable academic use.

Advanced Decoding and Prompting

Advanced decoding and stimulus techniques are effective ways of reducing hallucinations in LLMs and complement methods such as RAG and refinement (Tonmoy et al., 2024). Careful prompt engineering, using clear and

Cross Ref DOI: https://dx.doi.org/10.22161/ijtle.4.6.2

structured instructions, helps to constrain output, while Chain-of-Thought pushing increases reasoning by requiring step-by-step explanations, increasing the accuracy of complex tasks by up to 35% (Cossio, 2025). In educational contexts, it is particularly useful for problem solving, logical reasoning and structured feedback. Language-contrastive decoding (LCD) helps to further improve the accuracy of multimodal systems by combining outputs with text explanations and reducing object hallucinations (Manevich & Tsarfaty, 2024). The prompting of specific fields has also reduced errors in scientific tasks, including chemical hallucinations (Dahl et al., 2024), which highlights the importance of STEM education. In the future, hybrid systems combined with warnings, RAG, and human feedback may reduce hallucinations by 96%, while emerging neuro-symbolic architectures offer auditable and verified results for reliable educational applications.

Fine-Tuning and Model Alignment

Fine-tuning is a common strategy for aligning LLMs to specific domains or tasks. It involves retraining pre-trained models on small, high-quality datasets that are tailored to target applications, strengthening correct behaviour, and reducing hallucinations (Parthasarathy et al., 2024). For example, fine-tuning a curated legal dataset can improve the model's ability to cite precise legal provisions and produce reliable content. Despite its effectiveness, fine-tuning introduces a compromise between specialization and generalisation. Studies show that although specific domain adjustments improve task accuracy, they weaken the broader rationality and adaptation ability of models in unknown contexts (ZHAO et al., 2023). For example, an LLM finely adapted to a biology curriculum can show a reduction in performance when it comes to unrelated issues such as history. This limitation highlights that fine-tuning is not a universal remedy for hallucinations, but rather a strategy that depends on context and requires careful alignment with the intended use of the model.

Pedagogical and Curricular Interventions

The hallucination mitigation process is not just a technical effort; it also requires educational intervention. Teachers' practices and curriculum can be adapted to equip learners with the necessary skills to engage critically in AI-generated content (George, 2023). By promoting AI literacy, fact-checking habits, and critical thinking, students are better equipped to navigate the information ecosystem in which inaccuracies are presented easily and confidently. For example, the curriculum may include assignments in which students require LLMs to answer specific domain questions and then validate outputs against a trusted source. Such exercises not only show the strengths and weaknesses of the learners' AI systems but also promote epistemic vigilance. Similarly, the integration of modules on the use of responsible AI, such as appropriate citation practices, transparency in the reporting of AI support and strategies for ensuring claims, further strengthens academic integrity (Eze, 2024). These interventions highlight the human approach that complements the technical solutions and builds a resilient and informed user community capable of solving LLM hallucinations.

Developing Critical AI Literacy

Improving students' critical AI literacy is essential for effective hallucination mitigation. This literacy focuses on helping students develop "healthy skepticism" towards AI-generated content and advising them to approach outcomes critically. Students should be urged to assess the evidence, double-check assertions, and recognise the limitations of generative models rather than taking answers at face value. By having students critically evaluate confident but flawed LLM responses, compare them to verified references, and consider the dangers of misplaced trust, practical exercises can help students strengthen these abilities. These activities improve fact-checking skills and increase understanding of how LLMs create knowledge. The main objective is to prepare students to serve as responsible producers and astute consumers of AI-mediated information, while acknowledging that output fluency and confidence do not imply accuracy or dependability.

Redesigning Assessments and Assignments

Addressing the challenges of LLM hallucinations requires rethinking how we design assessments. Traditional fact-based tasks are increasingly easy to outsource to generative AI, which risks weakening authentic learning. To avoid this, educators should shift assessments toward higher-order thinking skills (synthesis, evaluation, and application) rather than simple recall. One effective approach is to require students to show evidence of their writing process through edit histories, draft submissions, or reflective commentaries, which places the focus on the intellectual journey rather than the final product. Complementing this, oral defences or verbal explanations of submitted work can push learners to engage more deeply with course material while making it harder to depend on AI tools. Together, these practices not only discourage academic dishonesty but also help students cultivate transferable skills that extend well beyond the classroom.

Cross Ref DOI: https://dx.doi.org/10.22161/ijtle.4.6.2

Dynamic Course Content Integration (DCCI)

Dynamic Course Content Integration (DCCI) allows for minimising hallucinations by incorporating verified Learning Management System (LMS) material into the LLM assistant, ensuring precise and contextual responses (Mzwri & Turcsányi-Szabo, 2025). For example, when students want to know the content of a conference, the system directly extracts slides or transcripts from the LMS. The challenges of interface design and privacy protection are still present, but DCCI provides a promising path to reliable AI-supported education.

CONCLUSION

This study examined hallucinations in LLMs with a focus on the implications for education. We stressed that hallucinations are structural effects of data limitations, model architecture, and prompting issues, not random errors. In educational contexts, these issues manifest misinformation, deterioration of academic integrity, unfair evaluation practices, and a decline in critical thinking. To mitigate these risks, we described pedagogical interventions including AI literacy programs, redesign of assessments, human supervision, and technical strategies, including retrieval-additional generation, advanced prompting, fine-tuning, and knowledge integration. The main contribution to this paper is its education-centred perspective, which links the general problem of hallucinations with the dangers facing students, teachers, and institutions. We believe that mitigation requires not only technical solutions, but also the integration of LLM into curriculum design, ethical frameworks and responsible teaching methods. To ensure fair adoption, future studies should evaluate mitigation techniques in different fields, languages, and socio-economic contexts. In the end, LLM should not replace but support human educators. These are strong helpers whose results are crucially filtered by human judgment. To build trust and ensure that AI-powered education improves rather than diminishes learning, this balance must be struck.

REFERENCES

[1] Attard, A., & Dingli, A. (2024). Empowering educators: Leveraging large language models to streamline content creation in education. ICERI2024 Proceedings,

[2] Bhattacharya, R. (2024). Strategies to mitigate hallucinations in large language models. *Applied Marketing Analytics*, *10*(1), 62-67.

[3] Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. Conference on fairness, accountability and transparency,

[4] Cherednichenko, O., Yanholenko, O., Badan, A., Onishchenko, N., & Akopiants, N. (2024). Large language models for foreign language acquisition. Proceedings of the CLW-2024: Computational Linguistics Workshop at 8th International Conference on Computational Linguistics and Intelligent Systems (CoLInS-2024), Lviv, Ukraine,

[5] Cossio, M. (2025). A comprehensive taxonomy of hallucinations in Large Language Models. *arXiv preprint arXiv:2508.01781*.

[6] Dahl, M., Magesh, V., Suzgun, M., & Ho, D. E. (2024). Large legal fictions: Profiling legal hallucinations in large language models. *Journal of Legal Analysis*, *16*(1), 64-93.

[7] Elsayed, H. (2024). The impact of hallucinated information in large language models on student learning outcomes: A critical examination of misinformation risks in AI-assisted education. *Northern Reviews on Algorithmic Research, Theoretical Computation, and Complexity*, *9*(8), 11-23.

[8] Eze, C. A. (2024). The role of educators in upholding academic integrity in an ai-driven era. *AI and ethics, academic integrity and the future of quality assurance in higher education*.

[9] Fang, X., Xu, W., Tan, F. A., Zhang, J., Hu, Z., Qi, Y., Nickleach, S., Socolinsky, D., Sengamedu, S., & Faloutsos, C. (2024). Large Language Models (LLMs) on Tabular Data: Prediction, Generation, and Understanding--A Survey. *arXiv preprint arXiv:2402.17944*.

[10] Feng, Y., Hu, H., Hou, X., Liu, S., Ying, S., Du, S., Hu, H., & Gao, Y. (2025). Hyper-RAG: Combating LLM Hallucinations using Hypergraph-Driven Retrieval-Augmented Generation. *arXiv preprint arXiv:2504.08758*.

[11] George, A. S. (2023). Preparing students for an AI-driven world: Rethinking curriculum and pedagogy in the age of artificial intelligence. *Partners Universal Innovative Research Publication*, *1*(2), 112-136.

[12] Gouscos, D. EJEL Volume 10 Issue 2.

[13] He, J., Rungta, M., Koleczek, D., Sekhon, A., Wang, F. X., & Hasan, S. (2024). Does prompt formatting have any impact on llm performance? *arXiv preprint arXiv:2411.10541*.

[14] Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM computing surveys*, *55*(12), 1-38.

[15] Lopez-Gazpio, I. (2025). Integrating Large Language Models into Accessible and Inclusive Education: Access Democratization and Individualized Learning Enhancement Supported by Generative Artificial Intelligence. *Information*, *16*(6), 473.

[16] Madsen, M., Lunde, I. M., Piattoeva, N., & Karseth, B. (2025). Introduction: Time and temporality in the datafied governance of education. *Critical Studies in Education*, *66*(2), 109-125.

[17] Manevich, A., & Tsarfaty, R. (2024). Mitigating hallucinations in large vision-language models (LVLMs) via language-contrastive decoding (LCD). *arXiv preprint arXiv:2408.04664*.

[18] Mirchandani, S., Xia, F., Florence, P., Ichter, B., Driess, D., Arenas, M. G., Rao, K., Sadigh, D., & Zeng, A. (2023). Large language models as general pattern machines. *arXiv preprint arXiv:2307.04721*.

[19] Mohammed, M., Al Dallal, A., Emad, M., Emran, A. Q., & Al Qaidoom, M. (2024). A comparative analysis of artificial hallucinations in GPT-3.5 and GPT-4: Insights into AI progress and challenges. *Business Sustainability with Artificial Intelligence (AI): Challenges and Opportunities: Volume 2*, 197-203.

[20] Mzwri, K., & Turcsányi-Szabo, M. (2025). Bridging LMS and Generative AI: Dynamic Course Content Integration (DCCI) for Connecting LLMs to Course Content--The Ask ME Assistant. *arXiv preprint arXiv:2504.03966*.

[21] Naser, M. (2025). From failure to fusion: A survey on learning from bad machine learning models. *Information Fusion*, *120*, 103122.

[22] Nazi, Z. A., & Peng, W. (2024). Large language models in healthcare and medical domain: A review. Informatics,

[23] Orenstrakh, M. S., Karnalim, O., Suarez, C. A., & Liut, M. (2024). Detecting LLM-generated text in computing education: Comparative study for ChatGPT cases. 2024 IEEE 48th Annual Computers, Software, and Applications Conference (COMPSAC),

[24] Parthasarathy, V. B., Zafar, A., Khan, A., & Shahid, A. (2024). The ultimate guide to fine-tuning llms from basics to breakthroughs: An exhaustive review of technologies, research, best practices, applied research challenges and opportunities. *arXiv preprint arXiv:2408.13296*.

[25] Perkins, M. (2023). Academic Integrity considerations of AI Large Language Models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching and Learning Practice*, *20*(2), 1-24.

[26] Pirjan, A., & PETROŞANU, D.-M. (2024). EXPLORING LARGE LANGUAGE MODELS IN THE EDUCATION PROCESS WITH A VIEW TOWARDS TRANSFORMING PERSONALIZED LEARNING. *Journal of Information Systems & Operations Management*, *18*(2).

[27] Razafinirina, M. A., Dimbisoa, W. G., & Mahatody, T. (2024). Pedagogical alignment of large language models (llm) for personalized learning: a survey, trends and challenges. *Journal of Intelligent Learning Systems and Applications*, *16*(4), 448-480.

[28] Selwyn, N. (2019). *Should robots replace teachers?: AI and the future of education*. John Wiley & Sons.

[29] Shahzad, T., Mazhar, T., Tariq, M. U., Ahmad, W., Ouahada, K., & Hamam, H. (2025). A comprehensive review of large language models: issues and solutions in learning environments. *Discover Sustainability*, *6*(1), 27.

[30] Sharma, S., Mittal, P., Kumar, M., & Bhardwaj, V. (2025). The role of large language models in personalized learning: a systematic review of educational impact. *Discover Sustainability*, *6*(1), 1-24.

[31] Toney, A. (2024). What Kind of Sourcery is This? Evaluating GPT-4's Performance on Linking Scientific Fact to Citations. Proceedings of the 1st Workshop on Customizable NLP: Progress and Challenges in Customizing NLP for a Domain, Application, Group, or Individual (CustomNLP4U),

[32] Tonmoy, S., Zaman, S., Jain, V., Rani, A., Rawte, V., Chadha, A., & Das, A. (2024). A comprehensive survey of hallucination mitigation techniques in large language models. *arXiv preprint arXiv:2401.01313*, *6*.

[33] ZHAO, X., LU, J., DENG, C., ZHENG, C., WANG, J., CHOWDHURY, T., YUN, L., CUI, H., XUCHAO, Z., & Zhao, T. (2023). Beyond one-model-fits-all: A survey of domain specialization for large language models. *arXiv preprint arXiv*, *2305*.