ISSN : 2455-9679
SJIF Impact Factor : 6.190

IJRTSM
INTERNATIONAL JOURNAL OF RECENT TECHNOLOGY SCIENCE & MANAGEMENT

"GENERATIVE AI FOR ACADEMIC RESEARCH AND LEARNING: A SYSTEMATIC COMPARATIVE REVIEW OF CHATGPT, GEMINI, GITHUB COPILOT, AND NOTEBOOKLM"

Shubhangi Patil
ATSS College Of Business Studies and Computer Application, Chinchwad, Pune

ABSTRACT

Generative Artificial Intelligence (GenAI) has rapidly transformed academic research and learning by providing intelligent support for content generation, knowledge synthesis, coding assistance, and personalized education. Recent AI tools such as ChatGPT, Gemini, GitHub Copilot, and NotebookLM have gained significant attention due to their ability to improve productivity and enhance learning experiences. These tools assist students, educators, and researchers in tasks including literature review, concept explanation, research summarization, academic writing, programming, and problem solving. Despite these advantages, concerns related to hallucination, academic integrity, plagiarism, privacy, and ethical misuse remain critical challenges. This review paper presents a comparative analysis of ChatGPT, Gemini, GitHub Copilot, and NotebookLM in the context of academic research and learning. The study evaluates these tools based on research assistance, learning support, coding capability, reliability, usability, and limitations. The analysis indicates that each tool offers distinct strengths depending on task requirements. ChatGPT performs strongly in reasoning and writing support, Gemini provides multimodal and web-assisted intelligence, GitHub Copilot excels in programming tasks, and NotebookLM demonstrates strong potential in source-grounded research synthesis. The findings suggest that generative AI can significantly enhance research productivity and learning effectiveness when used responsibly with critical human oversight. The study concludes that generative AI is likely to become an important collaborative technology in future academic ecosystems.

Keywords: *Generative Artificial Intelligence, ChatGPT, Gemini, GitHub Copilot, NotebookLM, Academic Research, Learning, Large Language Models.*

I. INTRODUCTION

Artificial Intelligence (AI) has become one of the most influential technological advancements in modern society, creating significant changes in education, research, healthcare, and software development. In recent years, the emergence of Generative Artificial Intelligence (GenAI) has further accelerated this transformation by enabling machines to generate human-like text, code, summaries, and analytical content. These capabilities are primarily driven by Large Language Models (LLMs) trained on massive datasets to understand and generate contextually meaningful responses. Unlike traditional AI systems that mainly perform classification or prediction tasks, generative AI systems support content creation and knowledge synthesis, making them highly relevant for academic and research applications [1].

The public release of ChatGPT by OpenAI in 2022 marked a major milestone in the adoption of generative AI tools for learning and research. The GPT-4 technical report demonstrated that large multimodal models can perform complex reasoning, problem-solving, summarization, and language generation tasks with high effectiveness across various professional and academic benchmarks [1]. This advancement led to widespread adoption of AI-based tools among

ISSN : 2455-9679
SJIF Impact Factor : 6.190

students, educators, and researchers for tasks such as literature exploration, note generation, writing assistance, and concept clarification. Further studies also indicate that advanced language models exhibit capabilities extending beyond text generation, including logical reasoning, coding support, and domain-specific problem solving [2].

Following the success of ChatGPT, several advanced generative AI tools entered the academic ecosystem, including Google Gemini, GitHubCopilot, and NotebookLM. Although all these tools are based on modern AI architectures, their core strengths differ according to application domains. ChatGPT and Gemini provide broad conversational and knowledge-assistance capabilities. GitHub Copilot specializes in programming support through AI-assisted code completion and debugging, while NotebookLM focuses on source-grounded research assistance by helping users interact with uploaded documents, notes, and research papers. This diversity has expanded the role of generative AI from general question answering to specialized academic workflows.

In academic research, generative AI tools are increasingly used to accelerate repetitive and cognitively demanding tasks such as literature review, summarization of scholarly articles, research gap identification, idea generation, code development, and manuscript refinement. Similarly, in educational settings, these tools support personalized learning by offering instant explanations, adaptive tutoring, and interactive feedback. Such capabilities can improve student engagement and reduce barriers in self-directed learning environments [3]. In computer science education and software engineering, AI-assisted programming tools have shown promising results in improving coding efficiency. Peng *et al.* reported that developers using GitHub Copilot completed programming tasks 55.8% faster compared to those without AI assistance, demonstrating the practical productivity benefits of AI-supported development environments [4].

Despite these advantages, the increasing dependence on generative AI introduces important academic and ethical challenges. One of the most critical concerns is hallucination, where AI systems generate convincing but factually incorrect information, including fabricated references, misleading explanations, or inaccurate conclusions. Such issues are especially problematic in academic research, where reliability and citation accuracy are essential. Additional concerns include plagiarism, overreliance on AI-generated content, reduced critical thinking, algorithmic bias, privacy risks, and academic integrity violations [5]. Therefore, while generative AI offers substantial benefits, responsible and critical usage remains necessary.

Although existing studies discuss the educational impact of generative AI, most research focuses primarily on individual tools, especially ChatGPT. Comparative studies involving multiple modern tools such as ChatGPT, Gemini, GitHub Copilot, and NotebookLM remain limited. Since each tool offers unique capabilities for research and learning, a systematic comparison is necessary to evaluate their strengths, limitations, and suitability for academic tasks.

This review paper presents a comparative analysis of ChatGPT, Gemini, GitHub Copilot, and NotebookLM in the context of academic research and learning. The study examines their major features, applications, benefits, limitations, and future potential. The findings are expected to help students, educators, and researchers select appropriate generative AI tools for enhancing learning outcomes and research productivity while ensuring ethical and responsible AI adoption.

II. LITERATURE REVIEW

Recent advancements in Generative Artificial Intelligence (GenAI) have significantly influenced academic research and learning environments. Large Language Models (LLMs) such as ChatGPT, Gemini, GitHub Copilot, and NotebookLM are increasingly being used for content generation, knowledge synthesis, coding assistance, and personalized tutoring. Existing literature highlights both the opportunities and challenges associated with integrating these tools into education and research workflows.

OpenAI's GPT-4 Technical Report established that large multimodal models can perform complex reasoning, summarization, and problem-solving tasks with strong performance across professional benchmarks [1]. This capability made ChatGPT one of the most widely adopted AI tools in academia for literature summarization, idea generation, academic writing assistance, and concept clarification. However, despite its effectiveness, concerns regarding hallucination and factual reliability remain significant.

Kasneci et al. analyzed the role of ChatGPT in education and reported that generative AI can enhance personalized learning, adaptive tutoring, and student engagement by providing instant feedback and explanations [3]. Their study also emphasized major concerns including overreliance on AI, academic dishonesty, reduced critical thinking, and

ISSN : 2455-9679
SJIF Impact Factor : 6.190

ethical misuse. The authors suggested that responsible AI integration requires human supervision and institutional guidelines.

Tzirides *et al.* explored the broader educational impact of generative AI and reported that AI-based systems can support assessment, feedback generation, and interactive learning [6]. Their findings indicate that generative AI can reduce repetitive instructional workload for educators while improving accessibility to learning resources. However, the study also highlighted limitations related to transparency, interpretability, and bias in AI-generated outputs.

In programming and software engineering education, GitHub Copilot has shown measurable productivity benefits. Peng *et al.* conducted an empirical study and found that developers using Copilot completed programming tasks 55.8% faster compared to non-users [4]. This suggests that AI-assisted coding tools can improve efficiency in code generation, debugging, and development workflows. Nevertheless, the study noted that generated code still requires human verification to ensure correctness, security, and maintainability.

Comparative evaluation of multiple generative AI tools remains limited. Alpar conducted a comparative study of ChatGPT-4, Google Gemini, and Microsoft Copilot in educational writing tasks and observed that all three tools performed effectively in brainstorming, outlining, and feedback generation, with no major statistical difference in most evaluation criteria [7]. The study reported that ChatGPT produced more actionable feedback, while Gemini showed strengths in idea clarity and diversity. This indicates that tool effectiveness varies depending on task requirements rather than overall model superiority.

Among emerging research-oriented AI tools, NotebookLM has gained attention due to its source-grounded approach based on Retrieval-Augmented Generation (RAG). Tufino examined NotebookLM as a collaborative learning tool and reported that its document-grounded responses improve traceability and reduce hallucination compared to conventional chatbots [8]. By generating responses from user-provided sources, NotebookLM supports evidence-based learning and literature synthesis, making it particularly useful for academic research.

Wagner discussed the application of generative AI in literature reviews and concluded that GenAI tools can significantly accelerate article screening, summarization, and knowledge extraction [12]. However, the study emphasized that AI-generated outputs cannot replace human judgment in critical analysis, synthesis, and research interpretation.

From the reviewed literature, it is evident that generative AI tools offer substantial benefits in academic research and learning through automation, personalization, and productivity enhancement. However, most existing studies focus on individual tools, particularly ChatGPT. Limited research provides a comprehensive comparison of ChatGPT, Gemini, GitHub Copilot, and NotebookLM within a unified academic framework. This gap motivates the present study, which aims to comparatively evaluate these tools based on their features, research utility, learning support, strengths, limitations, and future potential.

Author	Year	Tool Studied	Method	Key Findings	Research Gap
OpenAI	2023	ChatGPT / GPT-4	Technical Report	Strong reasoning, summarization, writing support	No academic comparison
Kasneci et al.	2023	ChatGPT	Educational Review	Useful in learning, tutoring	Integrity concerns
Peng et al.	2023	GitHub Copilot	Experimental	55.8% productivity increase	Limited to coding
Alpar	2025	ChatGPT, Gemini, Copilot	Comparative Study	All useful for academic writing	NotebookLM missing
Tufino	2025	NotebookLM	RAG-based analysis	Good source grounding	Limited comparative evidence

ISSN : 2455-9679
SJIF Impact Factor : 6.190

Although existing studies have evaluated individual generative AI tools in education and research, most investigations focus primarily on ChatGPT. Limited studies provide a comprehensive comparative analysis of ChatGPT, Gemini, GitHub Copilot, and NotebookLM within a unified academic framework. In particular, NotebookLM remains underexplored despite its growing importance in source-grounded research assistance. Therefore, a systematic comparative review is needed to evaluate these tools based on academic research support, learning effectiveness, strengths, limitations, and ethical considerations.

III. METHODOLOGY

This study adopts a comparative review methodology to analyze the impact of generative AI tools on academic research and learning. The review focuses on four widely used tools: ChatGPT, Gemini, GitHub Copilot, and NotebookLM. The objective is to compare these tools based on their capabilities, usefulness in academic tasks, strengths, limitations, and overall impact on research productivity and learning effectiveness.

A. Literature Selection Strategy

Relevant literature was collected from reliable academic databases and trusted sources, including Google Scholar, IEEE Xplore, Springer, Scopus-indexed journals, Elsevier, and official technical reports. The literature search was conducted using combinations of keywords such as:

1. “Generative AI in education”
2. “ChatGPT in academic research”
3. “Gemini AI for learning”
4. “GitHub Copilot productivity”
5. “NotebookLM research support”
6. “Large Language Models in higher education”

Only recent and relevant studies published between 2023 and 2026 were considered, as most generative AI tools gained significant adoption after 2022.

B. Inclusion and Exclusion Criteria

To ensure quality and relevance, the following criteria were applied.

Inclusion Criteria

1. Peer-reviewed journal articles, conference papers, and technical reports
2. Studies related to generative AI in education, learning, or research
3. Papers discussing ChatGPT, Gemini, GitHub Copilot, or NotebookLM
4. Studies published in English
5. Studies with clear findings related to academic applications

Exclusion Criteria

1. Blog posts, non-technical articles, and opinion-only content
2. Duplicate studies
3. Papers unrelated to academic research or learning
4. Studies lacking sufficient methodological details

After screening and filtering, the most relevant studies were selected for comparative analysis.

ISSN : 2455-9679
SJIF Impact Factor : 6.190

C. Comparison Criteria

The selected AI tools were evaluated using the following comparison parameters:

1) Research Assistance

Evaluation of the tool's ability to support:

1. literature review
2. paper summarization
3. idea generation
4. citation support
5. research gap identification

2) Learning Support

Assessment of how effectively the tool assists students in:

1. concept understanding
2. personalized tutoring
3. doubt solving
4. note generation
5. exam preparation

3) Coding and Technical Support

This criterion measures usefulness in:

1. code generation
2. debugging
3. code explanation
4. software development tasks

This parameter is especially important for GitHub Copilot.

4) Accuracy and Reliability

Evaluation of:

1. factual correctness
2. hallucination rate
3. response consistency
4. source grounding

Reliable output is critical in academic research.

5) User Experience and Accessibility

Comparison based on:

1. ease of use
2. interface quality
3. response speed

ISSN : 2455-9679
SJIF Impact Factor : 6.190

4. integration with external tools

D. Comparative Analysis Approach

Each selected tool was analyzed qualitatively using the above criteria. A comparative table was prepared to identify strengths and weaknesses of each tool. Based on the collected literature and feature analysis, the study evaluates which generative AI tools are most suitable for specific academic and learning tasks.

This methodology provides a structured framework for comparing modern generative AI tools and understanding their role in transforming academic research and education.

IV. COMPARATIVE ANALYSIS OF GENERATIVE AI TOOLS

This section presents a comparative analysis of four widely used generative AI tools—ChatGPT, Gemini, GitHub Copilot, and NotebookLM—with respect to their usefulness in academic research and learning. The comparison is based on research assistance, learning support, coding capability, reliability, and limitations.

A. ChatGPT

ChatGPT, developed by OpenAI, has emerged as one of the most widely adopted generative AI tools in education and research. Built on large language model architecture, it demonstrates strong capabilities in reasoning, summarization, explanation, and content generation. These capabilities make it useful for academic tasks such as literature review, concept clarification, brainstorming, drafting, and note preparation. The GPT-4 technical report highlighted strong performance in multiple professional and academic benchmark tasks, indicating its effectiveness in complex reasoning and problem-solving scenarios [1].

In educational settings, ChatGPT supports personalized learning by providing immediate explanations and interactive tutoring. However, an important limitation is hallucination, where the model may generate factually incorrect information or fabricated references. Therefore, outputs generated by ChatGPT require careful verification before academic use [3].

B. Gemini

Gemini, developed by Google, is a multimodal generative AI system designed to process and reason across text, images, audio, and other information formats. According to Google's technical report, Gemini demonstrates advanced reasoning across multimodal tasks and shows strong performance on language understanding benchmarks [101].

In academic contexts, Gemini can assist with concept explanation, information retrieval, summarization, and document analysis. Its integration with web-based information retrieval provides an advantage for accessing updated information. However, as with other generative AI tools, output quality depends on prompt design and source reliability. Peer-reviewed comparative research on Gemini in academic environments is still limited, indicating the need for further empirical evaluation.

C. GitHub Copilot

GitHub Copilot is an AI-powered coding assistant developed by GitHub in collaboration with OpenAI. Unlike general conversational AI tools, Copilot focuses specifically on software development tasks such as code completion, code generation, debugging, and programming assistance. This makes it highly relevant in computer science education and technical research.

Peng *et al.* conducted an empirical study on Copilot and reported that developers using the tool completed programming tasks 55.8% faster than those without AI assistance [4]. This suggests significant productivity improvement in coding workflows. However, AI-generated code may contain logical errors, insecure implementations, or inefficient solutions. Therefore, human validation remains essential before deployment in academic or industrial software projects [17].

D. NotebookLM

NotebookLM is a research-oriented AI tool developed by Google that emphasizes source-grounded interaction with user-provided documents. Unlike general-purpose chatbots, NotebookLM enables users to upload notes, PDFs, and

ISSN : 2455-9679
SJIF Impact Factor : 6.190

research articles, then generates responses based primarily on those sources. This source-aware design improves traceability and helps reduce hallucination risk.

Recent scholarly discussion suggests that retrieval-augmented generation (RAG)-based systems can improve evidence-based reasoning and support academic research more effectively than conventional generative AI systems [18]. As a result, NotebookLM shows strong potential for literature review, paper summarization, and research synthesis. However, because the tool is relatively new, peer-reviewed academic studies evaluating its educational impact remain limited.

E. Comparative Discussion

The comparative analysis indicates that each tool offers unique strengths depending on the academic task. ChatGPT performs strongly in reasoning, writing assistance, and interactive learning. Gemini provides multimodal understanding and access to updated web-based knowledge. GitHub Copilot is particularly effective for coding and software development tasks. NotebookLM offers strong support for source-based research and document-grounded analysis.

No single tool is universally optimal for all academic activities. The effectiveness of each tool depends on task requirements, user expertise, and the need for reliability. Therefore, students, educators, and researchers should select generative AI tools based on specific academic objectives while ensuring responsible and critical use of AI-generated outputs.

Criteria	ChatGPT	Gemini	Copilot	NotebookLM
Literature Review	High	High	Low	Very High
Concept Learning	High	High	Moderate	High
Coding Support	Moderate	Moderate	Very High	Low
Citation Reliability	Low–Moderate	Moderate	Low	High
Hallucination Risk	Moderate	Moderate	Low	Lower

V. CHALLENGES AND ETHICAL ISSUES

Despite the significant benefits of generative AI tools in academic research and learning, their increasing adoption introduces several technical, ethical, and educational challenges. These issues must be carefully addressed to ensure responsible and effective AI integration in academic environments.

A. Hallucination and Reliability Issues

One of the most critical limitations of generative AI systems is hallucination, where AI models generate outputs that appear convincing but contain factually incorrect, misleading, or fabricated information. In academic research, this becomes particularly problematic because inaccurate explanations or false references can negatively affect research quality and credibility. Large language models may produce fluent responses even when reliable evidence is unavailable, which can mislead students and researchers if outputs are accepted without verification [3]. Therefore, human validation remains essential when using AI-generated academic content.

B. Academic Integrity and Plagiarism

Generative AI has raised serious concerns regarding academic integrity, particularly in assignments, report writing, and research publication. Students may misuse AI tools to generate complete submissions without sufficient personal understanding or contribution. This creates challenges in evaluating genuine student learning and originality. Additionally, AI-generated content may unintentionally resemble training data patterns, creating concerns about plagiarism and authorship transparency. Institutions are therefore increasingly developing AI usage policies and ethical guidelines for responsible adoption [13].

C. Overdependence and Reduced Critical Thinking

ISSN : 2455-9679
SJIF Impact Factor : 6.190

Although AI tools improve productivity, excessive dependence on them may reduce independent thinking, creativity, and problem-solving ability. Continuous reliance on AI-generated explanations can weaken analytical reasoning and active learning processes. In educational settings, students may prioritize quick answers over conceptual understanding, reducing opportunities for deep learning and critical evaluation [2]. This challenge highlights the importance of using AI as an assistive tool rather than a replacement for human reasoning.

D. Bias and Fairness

Generative AI systems are trained on large-scale datasets collected from diverse sources. If these datasets contain social, cultural, or linguistic biases, the generated outputs may reflect or amplify those biases. Bias in AI-generated responses can affect fairness, inclusivity, and objectivity in educational and research contexts. Such bias may influence content recommendations, explanations, and decision support, potentially creating unequal learning experiences for different user groups [14]. Reducing bias remains a major research challenge in responsible AI development.

E. Privacy and Data Security

Privacy and security represent another major concern in the use of generative AI tools. Users often provide sensitive information such as research ideas, unpublished manuscripts, source code, institutional documents, or personal learning data while interacting with AI systems. If such information is improperly stored, processed, or exposed, it may create risks related to confidentiality, intellectual property, and data misuse. This concern is particularly important in academic research involving unpublished or proprietary work [15].

F. Ethical Use of AI in Research

The increasing use of AI in scholarly writing has initiated debates regarding authorship, accountability, and disclosure. AI systems can assist in drafting, summarization, language refinement, and literature synthesis, but they cannot assume responsibility for scientific accuracy, ethical judgment, or research interpretation. Major academic publishers and research organizations have clarified that AI tools cannot be listed as authors because they lack accountability and cannot take responsibility for published work [16]. Therefore, researchers must ensure transparent disclosure of AI usage and maintain full responsibility for the final scholarly output.

Overall, while generative AI tools provide substantial advantages in academic research and learning, their responsible use requires ethical awareness, institutional policies, and critical human oversight. Addressing these challenges is essential to maximize the benefits of AI while minimizing associated risks.

VI. FUTURE SCOPE

Generative Artificial Intelligence is evolving rapidly and is expected to play an increasingly important role in academic research and learning. Although current tools such as ChatGPT, Gemini, GitHub Copilot, and NotebookLM already provide substantial support for knowledge generation, coding assistance, and personalized learning, future advancements are likely to further improve their reliability, intelligence, and domain-specific applicability.

One important future direction is the development of more reliable and fact-grounded AI systems. Current generative AI models still suffer from hallucination and inconsistent citation generation. Future research is expected to focus on improving factual accuracy through techniques such as Retrieval-Augmented Generation (RAG), which combines large language models with external knowledge retrieval to generate evidence-based responses [9]. Such improvements can significantly enhance the trustworthiness of AI-generated outputs in academic research.

Another major area of future development is personalized and adaptive learning systems. Future AI tools are expected to move beyond general question answering toward intelligent tutoring systems that adapt according to individual learning style, performance, and pace. These systems may provide customized explanations, targeted practice tasks, and continuous feedback, enabling more effective student-centered learning environments [1]. This could improve accessibility and support learners with diverse educational backgrounds.

The integration of multimodal AI is also expected to expand significantly. Future generative AI systems will increasingly process and reason across text, images, audio, video, and structured documents within a unified

ISSN : 2455-9679
SJIF Impact Factor : 6.190

framework. Such capabilities can improve academic applications including visual learning, interactive simulations, research interpretation, and scientific communication. Multimodal AI may particularly benefit technical and scientific disciplines where diagrams, code, formulas, and visual content are essential for learning and analysis [10].

In software engineering and computer science education, AI-assisted programming tools are expected to become more sophisticated. Future coding assistants may support not only code generation but also software design recommendations, automated debugging, testing, security analysis, and architectural decision-making. This advancement can improve software development productivity and strengthen practical learning experiences in technical education [4].

Ethical and regulatory frameworks will also become increasingly important in the future. As generative AI becomes deeply integrated into academic workflows, institutions will need clear policies regarding AI-assisted writing, authorship disclosure, citation standards, privacy protection, and responsible usage. Future research should focus on establishing robust governance mechanisms that balance innovation with academic integrity and ethical responsibility [13].

Finally, more comparative empirical studies are needed to evaluate emerging AI tools in real academic settings. Current literature focuses heavily on individual tools, especially ChatGPT, while limited research compares multiple platforms such as ChatGPT, Gemini, GitHub Copilot, and NotebookLM using standardized evaluation metrics. Future work should include large-scale experimental studies involving students, educators, and researchers to measure learning outcomes, productivity gains, reliability, and long-term educational impact.

Overall, the future of generative AI in academic research and learning is highly promising. With improvements in reliability, personalization, multimodal intelligence, and ethical governance, generative AI has the potential to become a powerful collaborative partner in education and scholarly research rather than merely an assistive tool.

VII. CONCLUSION

Generative Artificial Intelligence has emerged as a transformative technology in academic research and learning. Tools such as ChatGPT, Gemini, GitHub Copilot, and NotebookLM are reshaping how students, educators, and researchers access knowledge, solve problems, generate content, and perform research-related tasks. These AI systems have demonstrated substantial potential in improving productivity, accelerating information processing, enhancing coding efficiency, and supporting personalized learning experiences.

The comparative analysis conducted in this study shows that each tool offers unique advantages in specific academic contexts. ChatGPT provides strong reasoning, explanation, and writing assistance; Gemini offers multimodal understanding and web-integrated support; GitHub Copilot significantly improves coding productivity in software-related tasks; and NotebookLM enables more reliable document-based research support through source-grounded interaction. Therefore, no single tool can be considered universally superior for all academic applications.

Despite these benefits, challenges such as hallucination, plagiarism, ethical misuse, privacy concerns, and overdependence remain major limitations. These issues emphasize the importance of critical thinking, human verification, and responsible AI adoption in academic environments. AI-generated outputs should support, rather than replace, human intelligence and scholarly judgment.

In conclusion, generative AI has the potential to become a powerful collaborative partner in education and research. With continued advancements in accuracy, multimodal intelligence, and ethical governance, these tools can significantly contribute to the future of higher education and scholarly innovation. However, their successful integration requires balanced use, clear institutional policies, and continuous human oversight to ensure reliability, originality, and academic integrity.

ISSN : 2455-9679
SJIF Impact Factor : 6.190

REFERENCES

[1] OpenAI, "GPT-4 Technical Report," arXiv preprint arXiv:2303.08774, 2023.

[2] S. Bubeck, V. Chandrasekaran, R. Eldan, et al., "Sparks of Artificial General Intelligence: Early Experiments with GPT-4," arXiv preprint arXiv:2303.12712, 2023.

[3] S. Kasneci, E. Sessler, M. Küchemann, et al., "ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education," Learning and Individual Differences, vol. 103, Art. 102274, 2023.

[4] S. Peng, E. Kalliamvakou, P. Cihon, and M. Demirer, "The Impact of AI on Developer Productivity: Evidence from GitHub Copilot," arXiv preprint arXiv:2302.06590, 2023.

[5] J. Gallifant, B. K. Steenhuisen, and D. S. Martin, "Peer Review of GPT-4 Technical Report and System Card," Cureus, vol. 16, no. 1, 2024.

[6] A. O. Tzirides et al., "Generative AI: Implications and Applications for Education," arXiv preprint arXiv:2305.07605, 2023.

[7] Ö. Alpar, "Evaluating Generative AI Tools for Improving English Writing Skills: A Preliminary Comparison of ChatGPT-4, Google Gemini, and Microsoft Copilot," European Journal of Educational Research, vol. 14, no. 4, pp. 1291–1308, 2025.

[8] E. Tufino, "NotebookLM: An LLM with RAG for Active Learning and Collaborative Tutoring," arXiv preprint arXiv:2504.09720, 2025.

[9] G. Wagner, "Generative Artificial Intelligence for Literature Reviews," Journal of Information Technology, 2026.

[10] Google, "Gemini: A Family of Highly Capable Multimodal Models," Technical Report, 2023.

[11] A. Mastropaolo et al., "An Empirical Study on the Robustness of AI Code Generators," 2023.

[12] P. Lewis et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," Advances in Neural Information Processing Systems (NeurIPS), 2020.

[13] UNESCO, Guidance for Generative AI in Education and Research, Paris, France: UNESCO, 2023.

[14] National Institute of Standards and Technology (NIST), AI Risk Management Framework (AI RMF 1.0), NIST AI 100-1, Gaithersburg, MD, USA, Jan. 2023.

[15] European Union Agency for Cybersecurity (ENISA), AI Cybersecurity Challenges, Heraklion, Greece: ENISA, 2024.

[16] Nature Editorial, "Tools such as ChatGPT threaten transparent science; here are our ground rules for their use," Nature, vol. 613, no. 7945, p. 612, 2023.